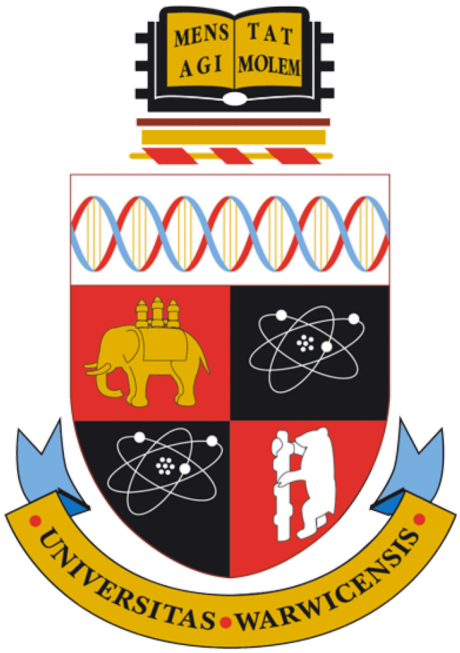
Mechanism of Pharmacophore Biosynthesis for Epoxyketone Proteasome Inhibitors

Joshua W. Cartwright

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Chemistry



University of Warwick

Department of Chemistry

September 2017

Contents

List of figures, tables and schemes

Acknowledgements

To J.W.

My star, my perfect silence.

Firstly, I'd like to thank supervisor Prof. Greg Challis for allowing me to work on such an interesting and varied project, and being an invaluable source of knowledge and guidance throughout my PhD, plus exhaustively going through my thesis.

I'd like to thank everyone who took the time to teach me the necessary skills for my research; for biology Dr. Dani Zabala patiently walked me through gene cloning, and Dr. Matthew Jenner helped a great deal with the ins and outs of protein purification. For chemistry, Dr. Doug Roberts, as well as Christian Hobson, Chris Perry and Rob Jenkins all imparted their chemical wisdom at various points throughout my PhD. For spectroscopy, Dr. Lijiang Song, Dr. Ivan Prokes as well as Dani and Matt again provided great assistance in obtaining nice spectra.

Thanks very much to Dr. Lona Alkalaf, who also provided guidance in the latter half of my PhD and painstakingly read through my early (and very sketchy) thesis drafts.

I'd like to thank Rakesh, Magda, Richard, Shanshan, Yousef and Jade for making the office a brighter place, particularly the latter two for putting up with my clutter! In the area I shared in the chemistry lab, Chris, Christian, Rob and Dan made work much more entertaining. Thanks in particular to Christian, for being a great fume hood partner, an appreciator of fine music, and putting up with both my sombre demeanour and occasional crazy rants! Thanks to the MIBTP/BBSRC for funding, my advisory panel, and to everyone in the Challis group and CBRF I haven't already mentioned.

Last but not least I'd like to thank Jess, for sticking by me through both the ups and downs as well as keeping me (mostly) sane. I couldn't have done it without you.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any other degree. Results from other authors are referenced in the usual manner throughout the text.

Parts of this thesis have been published by the author:

"A Flavin-Dependent Decarboxylase–Dehydrogenase–Monooxygenase Assembles the Warhead of α,β-Epoxyketone Proteasome Inhibitors" by Zabala, D.; Cartwright, J. W.; Roberts, D. M.; Law, B. J. C.; Song, L.; Samborskyy, M.; Leadlay, P. F.; Micklefield, J.; Challis, G. L. *J. Am. Chem. Soc.* **2016**, *138*, 4342.

Date: _____

Joshua W. Cartwright

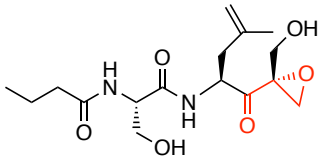
Abbreviations

A	Adenylation
ACAD	Acyl-CoA dehydrogenase
ACP	Acyl carrier protein
APS	Ammonium persulfate
AT	Acyl transferase
ATP	Adenosine triphosphate
BGC	Biosynthetic gene cluster
BOC	*tert*-Butyloxycarbonyl protecting group
BOP	(Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate
C	Condensation
CDI	Carbonyl diimidazole
cMT	Carbon-methyltranferase
CoA	Coenzyme A
CORM3	Carbon monoxide releasing molecule 3
DCC	*N,N'*-Dicyclohexylcarbodiimide
DCM	Dichloromethane
DHCA	Dicyclohexylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
FAD	Flavin adenine dinucleotide
FAS	Fatty acid synthase
FMN	Flavin mononucleotide
GC	Guanosine cytosine
HOBt	Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HR	High resolution
IC_{50}	Half maximal inhibitory concentration

IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KS	Ketosynthase
LB	Luria-Bertani
LC	Liquid chromatography
MALDI	Matrix assisted laser desorption/ionization
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
NADH	Reduced nicotinamide adenine dinucleotide
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance
NRPS	Nonribosomal peptide synthetase
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PKS	Polyketide synthase
PLE	Pig liver esterase
PMSF	Phenylmethylsulfonyl fluoride
PPT	Phosphopantethienyl
Q	Quadrupole
rRNA	Ribosomal ribonucleic acid
SAH	*S*-adenosyl homocysteine
SAM	*S*-adenosyl methionine
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
STE	Sodium chloride, Tris, EDTA buffer
TBE	Tris, boric acid, EDTA buffer
TBS	*tert*-Butyldimethylsilyl protecting group
TE	Thioesterase
TEMED	Tetramethylenediamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TOF	Time-of-flight
tRNA	Transfer ribonucleic acid
TSB	Tryptic soy broth
UHPLC	Ultra high performance liquid chromatography
UV	Ultra-violet

Abstract

Natural α,β-epoxyketone proteasome inhibitors produced primarily by *Streptomyces* species, such as TMC-86A, are potent anticancer compounds. The enzymes responsible for biosynthesis of their key α,β-epoxyketone pharmacophore had yet to be determined at the onset of this research. Elucidation of the responsible enzymes and their putative substrate would allow for stereoselective completion of a synthetically challenging epoxidation reaction, responsible for the high cost of epoxyketone pharmaceuticals produced by industry.



TMC-86A

Abstract figure: The α,β-epoxyketone proteasome inhibitor TMC-86A from *Streptomyces chromofuscus*. The epoxyketone pharmacophore is highlighted in red.

The cytochrome P450 TmcI and the flavin-dependent TmcF enzymes from the TMC-86A biosynthetic gene cluster in *Streptomyces chromofuscus*, believed to have a role in epoxyketone biosynthesis, were cloned, expressed and overproduced. Potential substrates were produced through a peptide coupling synthetic route. *In vitro* assays demonstrated that EpnF, a TmcF homologue, could produce α,β-epoxyketone compounds with an α-dimethyl-β-keto carboxylic acid substrate. These substrates were prone to decarboxylative degradation, therefore stable methyl ester precursors were synthesised as an alternative. Using an esterase, methyl ester precursors were hydrolysed to reproduce the previously observed epoxyketone compound in a one pot reaction with EpnF. Substrate specificity of EpnF was probed by *in vitro* assay using an alanine derived analogue, which was found to be a viable substrate for the enzyme. Lastly, the EpnF catalytic cycle was interrogated with a synthesised authentic standard of a proposed intermediate, which was observed to co-elute with the desired intermediate from a quenched EpnF assay using UHPLC-MS chromatography.

Chapter 1: Introduction

1.1 Natural products

Natural products produced by plants, bacteria and fungi have an extraordinary range of bioactivities, and can serve as vital leads in the development of pharmaceutically and agriculturally relevant compounds.[1,2] Anticancer activity is found in natural products such as taxol **1** (figure 1.1A), a terpene produced in the bark of *Taxus breuifolia*.[3–5] Many natural products exhibit antibacterial activity, for example the 26S rRNA binder erythromycin A **2** (figure 1.1B), a polyketide marcrolide produced by bacterium *Saccharopolyspora erythraea*.[6,7] For crop protection, the nonribosomal peptide bassianolide **3** (figure 1.1C) from the fungus *Beauveria bassiana* is an exemplar of a potent insecticidal virulence factor.[8] While bioactive and beneficial natural products are produced from an extensive range of sources, some of the most prolific producers of clinically and agriculturally relevant compounds are *Streptomyces* species.

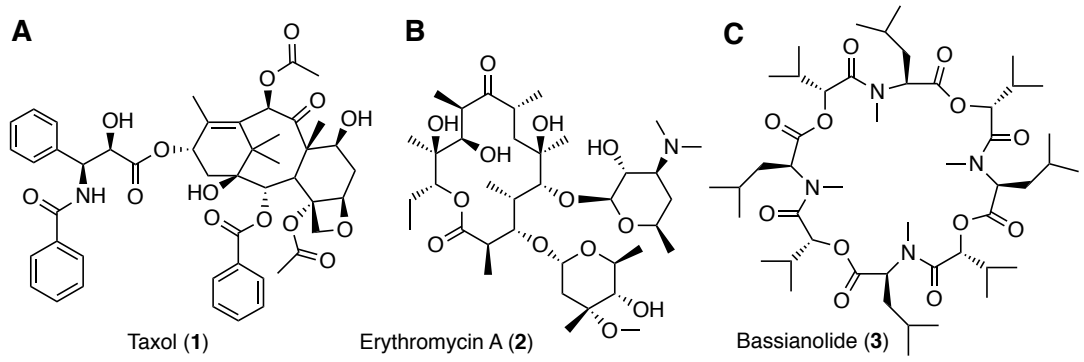


Figure 1.1. A: Structure of the terpene taxol from *Taxus breuifolia*. B: Structure of the macrolide erythromycin A from *Saccharopolyspora erythraea*. C: Structure of peptide bassianolide from *Beauveria bassiana*.

1.1.1 *Streptomyces* natural products

Streptomyces species are filamentous Gram-positive bacteria from the actinobacteria phylum, with high GC content DNA.[9] Their life-cycle revolves around the dispersal of spores from aerial hyphae, which follows their initial growth as a vegetative

mycelium through soil-bound branching hyphae.[10] *Streptomyces* are the focus of great academic and industrial interest for being the foremost producers of bioactive secondary metabolites (figure 1.2), with some models predicting production of over 100000 antimicrobial compounds from the genus as a whole of which contemporary research has uncovered but a fraction.[11]

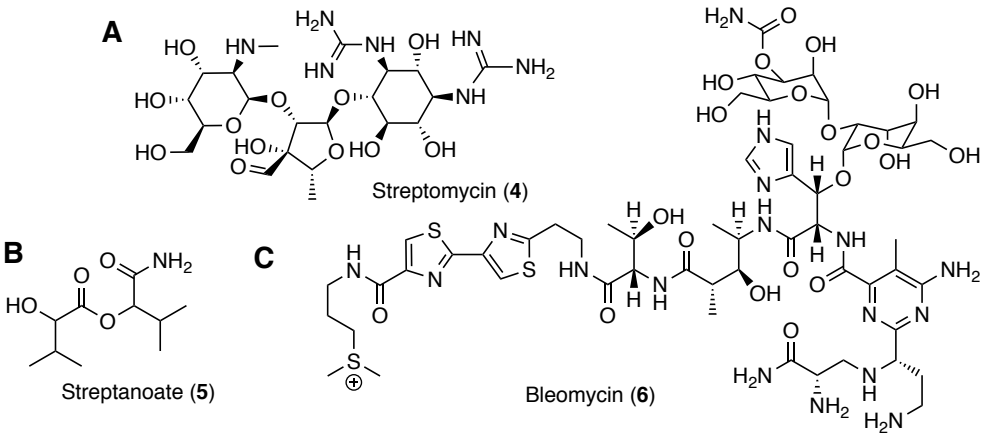


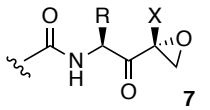
Figure 1.2. A: Structure of streptomycin. B: Structure of streptanoate. C: Structure of bleomycin.

One of the earliest discovered *Streptomyces* natural products was the trisaccharide streptomycin **4**, isolated from *Streptomyces griseus* in 1943 (figure 1.2A).[9] Streptomycin is a potent antibiotic still in use today, with discovery of its efficacy against *Mycobacterium tuberculosis* earning Selman Waksman the 1952 Nobel prize.[12,13] Anticancer compounds have also been routinely discovered in *Streptomyces* species. A recent example of one such compound is streptanoate **5** from *Streptomyces* sp. DE3 (figure 1.2B), putatively the first bioactive butanoate discovered in a *Streptomycete*.[14] Streptanoate was demonstrated to have a lower half maximal inhibitory concentration (IC_{50}) against some cancer cell lines than established anti-cancer pharmaceuticals, while having limited toxicity against healthy cell lines. A well-established clinically approved antitumoral natural product is the hybrid non-ribomsomal peptide synthetase (NRPS) polyketide synthase (PKS) derived bleomycin from *S. verticillus*.[15] Bleomycin **6** anticancer activity is due to its effectiveness at DNA

scission, and is used to treat a variety of cancers, though resistance to bleomycin continues to develop.[16] NRPS-PKS hybrid epoxyketone proteasome inhibitors, derived primarily from actinomycetes including *Streptomyces* species, are a further example of an anticancer class of natural products.

1.2 Epoxyketone compounds

Epoxyketone compounds are irreversible and highly selective inhibitors of the 26S proteasome.[17] Their pharmacophore is the terminal α,β-epoxyketone moiety **7**, comprised of adjacent electrophilic epoxide and ketone functionalities (figure 1.3).



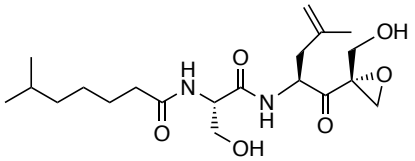
7

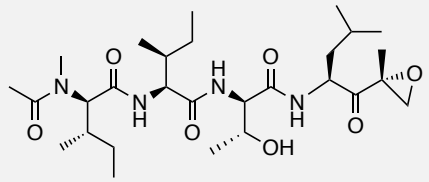
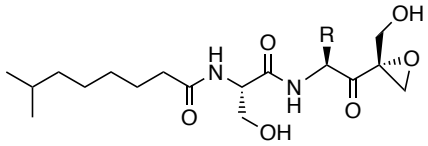
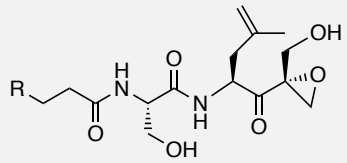
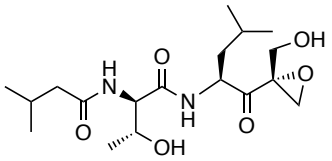
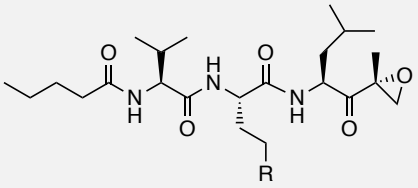
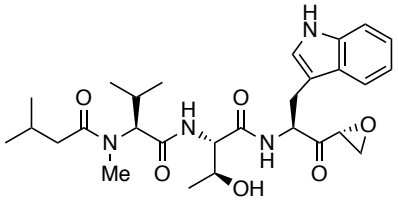
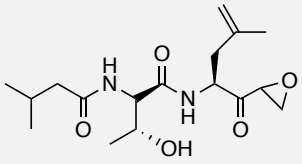
X = Me, CH_2OH or H

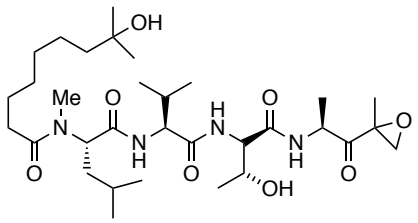
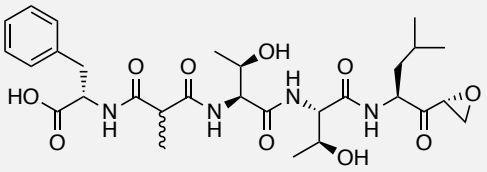
Figure 1.3. The α,β-epoxyketone pharmacophore.

The backbone of epoxyketone compounds typically incorporates polyketide, nonribosomal peptide and fatty acid elements. To date, a wide range of epoxyketone proteasome inhibitors have been reported in the literature (table 1.1).

Table 1.1: Reported epoxyketone proteasome inhibitors with stereochemistry where known, arranged chronologically, with predicted or confirmed stereochemistry.

Compound *species*	Structure	Year discovered
Eponemycin (**8**) *S. hygroscopicus*[18]		1990

Epoxomicin (**9**) *Actinomycetes* strain No. Q996-17[19]		1992
Epopromycin A (**10**)/B (**11**) *S.* sp. NK04000[20]	 R = iBu, Epopromycin A R = $CH_2C{=}CH_2(CH_3)$, Epopromycin B	1997
TMC-86A (**12**)/B (**13**) *S.* sp. TC 1084[21]	 R = Me, TMC-86A R = $CH_2CH_2COH(CH_3)_2$, TMC-86B	1999
TMC-96 (**14**) *Saccharothrix* sp. TC 1094[21]		1999
Carmaphycin A (**15**)/B (**16**) *Symploca* sp[22]	 R= $SOCH_3$, Carmaphycin A R = SO_2CH_3, Carmaphycin B	2012
Tryptopeptin A (**17**) *Streptomyces* sp. KUSC-G11[23]		2015
Landepoxcin A (**18**) soil metagenome library[24]		2015

Clarepoxcin A (**19**) soil metagenome library[24]		2015
Macyranone A (**20**) *Cystobacter fuscus*[25]		2015

While other epoxyketone compounds exist, they are not known to inhibit the 26S proteasome. Natural products such as asukamycin **22** from *S. nodosus* subsp. *asukaensis* and manumycin **21** from *S. parvulus* have cyclic epoxyketones located centrally in their respective structures which contribute to their antibacterial and antifungal activity rather than inhibiting the proteasome.[26,27] Trapoxin A **23** from *Helicoma ambiens* does have a terminal epoxyketone, but has been demonstrated to act as an irreversible inhibitor of histone deacetylase HDAC8.[28,29] While still possessing a mechanism of action dependent on nucleophilic attack by the biological target, presumably the peptide backbone of the other epoxyketones is important in producing specificity for the proteasome which trapoxin A lacks.

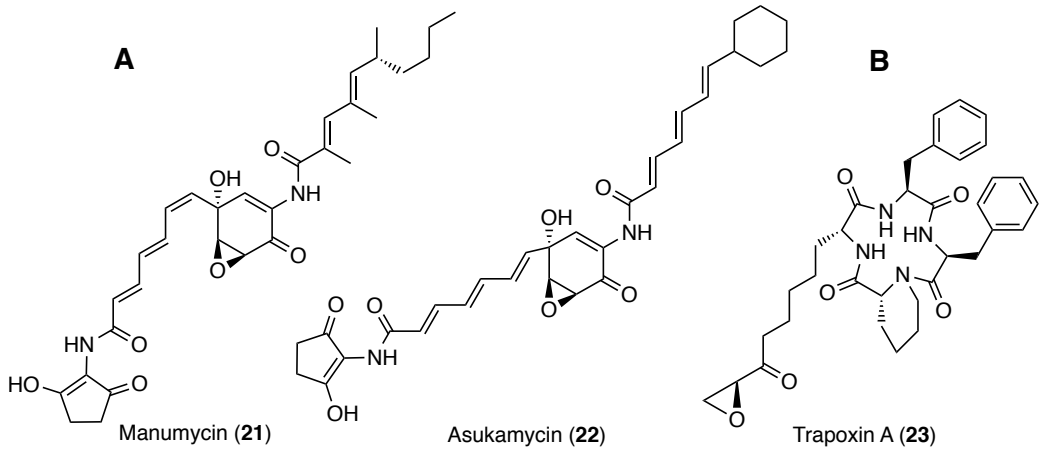


Figure 1.4. A: Structures of the cyclic epoxyketones manumycin and asukamycin. B: Structure of linear epoxyketone trapoxin A, a histone deacetylase HDAC8 inhibitor.

1.2.1 The 26S proteasome

The 26S proteasome is a threonine protease responsible for degradation of polyubiquitinated proteins, and is comprised of a cylindrical 20S core particle with 19S regulatory particles attached to each end.[30] Proteins are ubiquitinated by action of E1 (ubiquitin activating enzyme), E2 (ubiquitin carrier protein) and E3 enzymes (ubiquitin ligase) respectively, attaching the 76 amino acid peptide to a protein's lysine residue. Ubiquitination of position 48 lysine of ubiquitin that is already attached to a protein builds a polyubiquitin chain.[31]

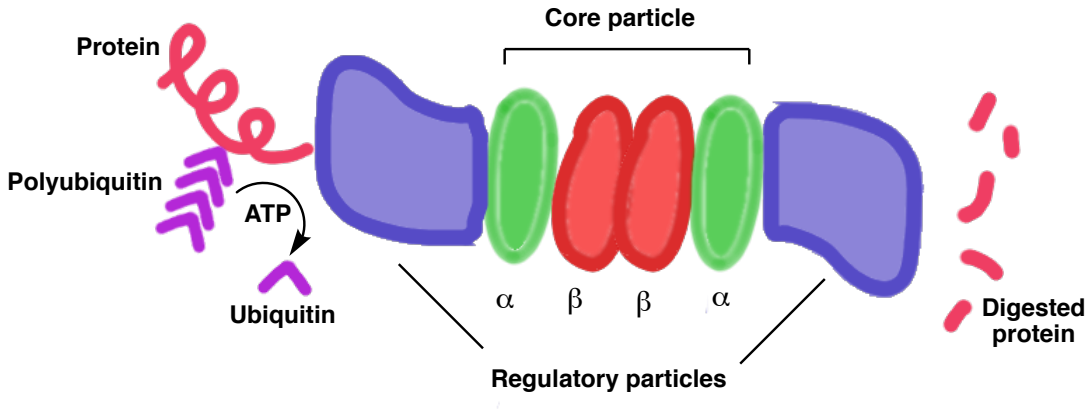


Figure 1.5: Structure of the 26S proteasome, demonstrating proteolysis of a polyubiquitinated protein.

The regulatory particle is composed of 19 subunits, which work together to recognize a polyubiquitinated protein substrate, cleave off the polyubiquitin chain, and then by action of a hexameric group of 6 ATPase subunits unfold and translocate the protein through a pore in the core particle.[30] The 20S core particle is made up of four heptameric subunits, a core of two β subunits where degradation of substrates takes place by action of *N*-terminal threonine residues, capped by two α subunits which maintain an aperture for protein translocation when in complex with the regulatory particle (figure 1.5).[32]

This degradation pathway, apart from assisting with cell homeostasis and providing a protein quality control mechanism, also has a role in transcriptional regulation. For

this reason cancers like multiple myeloma depend heavily on action of the proteasome, making epoxyketone compounds effective anticancer therapies.[17,31]

1.2.2 Mechanism of proteasome inhibition by epoxyketone compounds

Epoxyketone binding to the 26S proteasome was first discovered through feeding of a biotin labelled eponemycin analogue to murine thymoma cell line EL4. Denaturation of proteins in the cell lysate and their isolation using avidin-agarose, trypsin digest and subsequent matrix assisted laser desorption ionization-time of flight (MALDI-TOF) spectroscopy revealed the biotin tagged proteins to be subunits of the proteasome.[33]

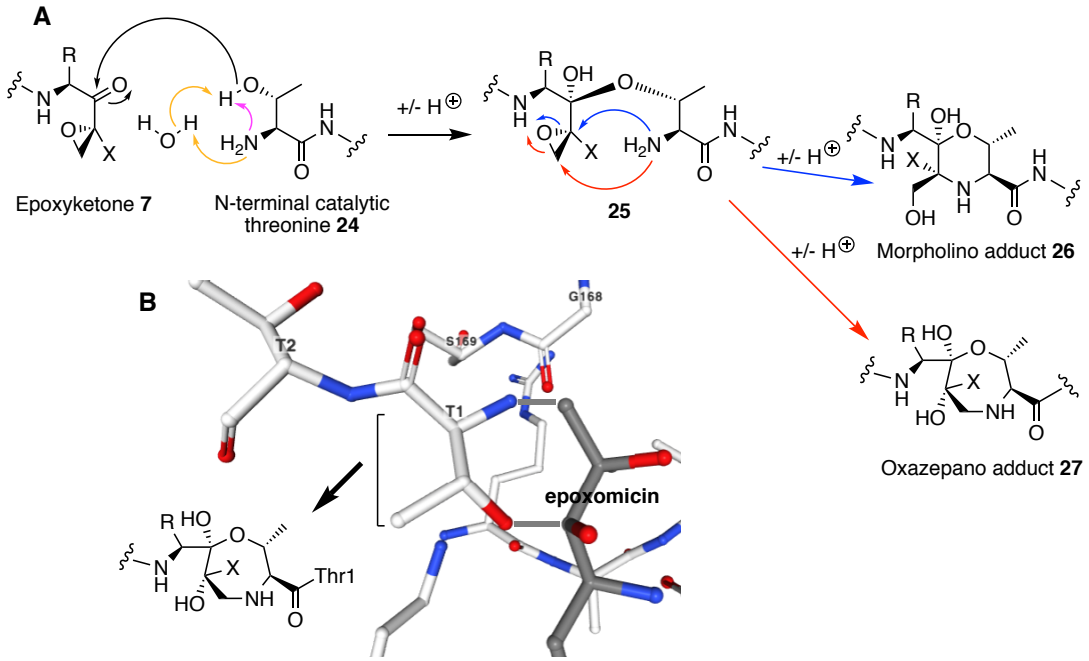


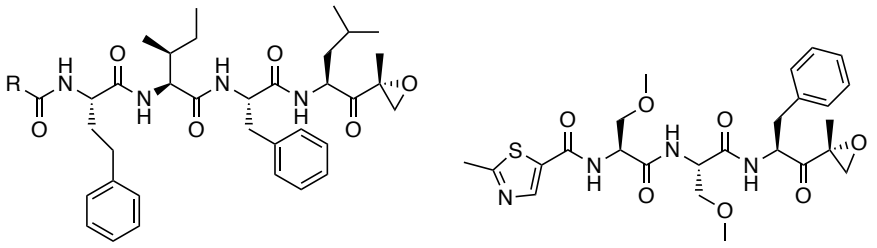
Figure 1.6. A: Proposed mechanism of action of epoxyketone proteasome inhibitors, demonstrating initial nucleophilic attack either by intermolecular or intramolecular activation, followed by 6 *Exo-tet* (top route) or the more recently proposed 7 *Exo-tet* (bottom route) ring closure. B: Active site of human 20S core particle crystallised with epoxomicin by Schrader and coworkers.[34]

The specific mechanism of action for inhibition of the 26S proteasome was elucidated with the crystal structure of 20S proteasome core particle bound to epoxomicin, suggesting covalent morpholine moiety **26** between epoxomicin and the *N*-terminal catalytic threonine of the proteasome. This strongly implicated a mechanism of nucleophilic attack on the electrophilic pharmacophore, resulting in 6 *Exo-tet* ring closure favoured by Baldwin's rules due to this angle of attack being achievable from an energetically favourable conformation (figure 1.6A).[35,36] More recent higher resolution crystallographic data with synthetic epoxyketone oprozomib indicated 7 *Exo-tet* ring closure with this threonine residue to produce oxazapan adduct **27** was more likely (figures 1.6A and 1.6B).[37,34] If accurate this revised mechanism is interesting, as aside from epoxyketones which have a second electrophile located at both the α and β positions relative to the ketone functionality, it proposes other more trivially synthesised β-keto electrophiles may be suitable proteasome inhibitors.

The proposal of a mechanism of action for proteasome inhibition, in addition to a high degree of specificity for the proteasome itself compared to existing boronic acid based proteasome inhibitors, lead to an increased interest in industry to develop synthetic epoxyketone compounds.

1.2.3 Epoxyketone pharmaceuticals and their synthesis

In the first attempt to epoxyketone pharmaceuticals, Crews *et al* investigated synthesis of an epoxyketone with improved activity over epoxomicin through a classic medicinal chemistry approach.[38] Initially the length of the peptide leading into the epoxyketone pharmacophore was varied, followed by variation of the amino acid side chains when optimal length had been determined. Finally by altering derivatization following the terminal amide functionality on the *N*-terminus, the potent proteasome inhibitor YU-101 was produced.



R = CH_3, YU-101 (**28**)
R = $CH_2N(CH_2CH_2)_2O$, Carfilzomib (**29**)

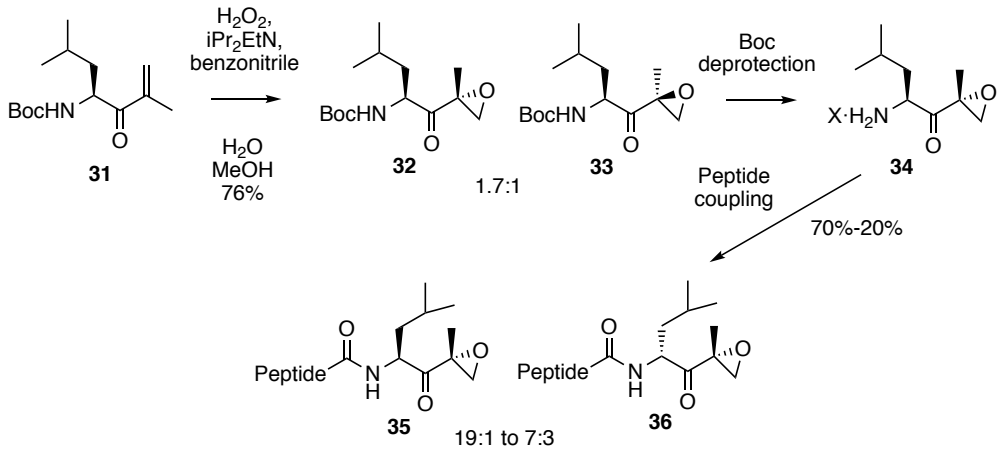
Oprozomib (**30**)

Figure 1.7: Structures of synthetic epoxyketones YU-101, carfilzomib and oprozomib.

With YU-101 in hand, the San Franciso-based company Proteolix set about trying to improve on the compound's efficacy, producing a library of hundreds of variants.[17] Through the screening process however, it was found that the bioactivity of YU-101 **28** could not be readily improved upon. The remaining issue with YU-101 **28** was its poor water solubility, and to rectify this a morpholine adduct was attached to the *N*-terminus, producing carfilzomib **29** (figure 1.7). Carfilzomib **29** is currently licenced to treat the cancer multiple myeloma, and is effective against cancer that is resistant to the previous generation of boronic acid inhibitors. A second generation of epoxyketone compounds, including orally available oprozomib **30**, are currently undergoing clinical trials (figure 1.7).[39]

While synthesis of the peptide portion of these synthetic epoxyketones proceeds through trivial solid phase peptide coupling using a coupling reagent (coupling reagents are discussed in sections 3.1.2.2, 3.1.2.3 and 4.1.3), the synthesis of the epoxyketone pharmacophore is more challenging. Using a Boc-protected α,β-unsaturated ketone derived from Weinreb-Nahm ketone synthesis (section 3.1.2.3), the routinely used basic epoxidation of the alkene bond using hydrogen peroxide, benzonitrile and Hunig's base is cited as resulting in a 76% yield with a stereoselectivity of 1.7:1 (63%:37%) for the desired enantiomer – with epimerisation resulting basic deprotonation of the acidic α-proton as well as epimerisation resulting from use of a coupling reagent (see section 3.1.2.1 and figure 3.3 for more detail).[40] After Boc-deprotection, subsequent coupling to the solid phase synthesised peptide

has been reported to result in an additional 5% to 30% epimerization of the stereocenter adjacent to the epoxyketone, with yields varying from 20-70%.[38]



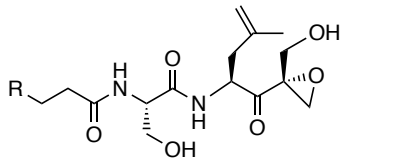
Scheme 1.1: Key losses in enantioselectivity during epoxyketone synthesis.

This synthetic process is far from enantioselective, resulting in high production costs and an estimated $10000 cost for a typical 28-day cycle of carfilzomib **29**. Discovery of an enzyme which could catalyse efficient epoxyketone formation in aqueous conditions and with a high degree of enantioselectivity, without the need for organic solvents or strongly oxidizing reagents like hydrogen peroxide, would help to lower the cost of industrial synthesis by increasing efficiency. In order to identify the enzyme responsible for pharmacophore formation in nature, the biosynthetic gene clusters of TMC-86A **12** and eponemycin **8** were investigated (table 1.1, figure 1.8).

1.3 TMC-86A and eponemycin

Before the onset of this project, production of TMC-86A **12** was discovered in *S. chromofuscus* ATCC49982 by the Challis group, and a biosynthetic gene cluster most likely to be responsible identified within the sequenced genome. Production of structurally related eponemycin **8** and its gene cluster from *S. hygroscopicus* ATCC53709 had already been reported by Moore *et al.*[41] As both compounds are relatively short for epoxyketones (figure 1.8), with just two additional amino acid

stereocenters in addition to the epoxyketone stereocenter, they were ideal candidates for the purposes of synthesising substrate analogues.



R = Me, TMC-86A (**12**)
R = $CH_2CH_2CH_2(CH_3)_2$, eponemycin (**8**)

Figure 1.8: Structures TMC-86A and eponemycin.

1.3.1 Biosynthetic gene cluster

Biosynthesis of TMC-86A **12** was linked to the *tmc* gene cluster through preparation of a *tmcH* knockout mutant in *S. chromofuscus* by Dr. D. Zabala, which abrogated production of TMC-86A **12**.[42] Similarly, whole pathway expression of the *epn* cluster in *S. albus* by Moore *et al* revealed eponemycin to be the major metabolite.[41] Both the *tmc* and *epn* clusters have the same arrangement of homologous genes (figure 1.9A), and so TMC-86A **12** and eponemycin **8** are predicted to form through analogous biosynthetic routes.

The proposed biosynthesis (figure 1.9B), covered in detail in the following sections, begins with condensation of an acetyl-CoA primer **37**, catalysed by the FabH-like fatty acid synthase (FAS) enzyme TmcD, to a malonyl unit bound to the phosphopantetheine **43** (PPT) modified acyl carrier protein (ACP) encoded by *tmcE*.[43] The resulting *β*-keto thioester then undergoes reduction by normal *Streptomyces* FAS metabolism to form butanoyl-ACP **38**. The two NRPS modules (groups of catalytic domains) encoded by *tmcG* ensure transpeptidation of the acyl chain by first L-serine followed by L-leucine.[44] Desaturation of the leucine side chain is proposed to be catalysed by the cytochrome P450 encoded by *tmcK*, with initial hydroxylation followed by dehydration to provide the alkene bond, though the timing of this tailoring step is unclear. Activity assays of the second adenylation (A) domain in the NRPS

TmcG demonstrated greater preference for L-leucine than 4,5-dehydro-L-leucine, implying that desaturation takes place following action of TmcG.[42]

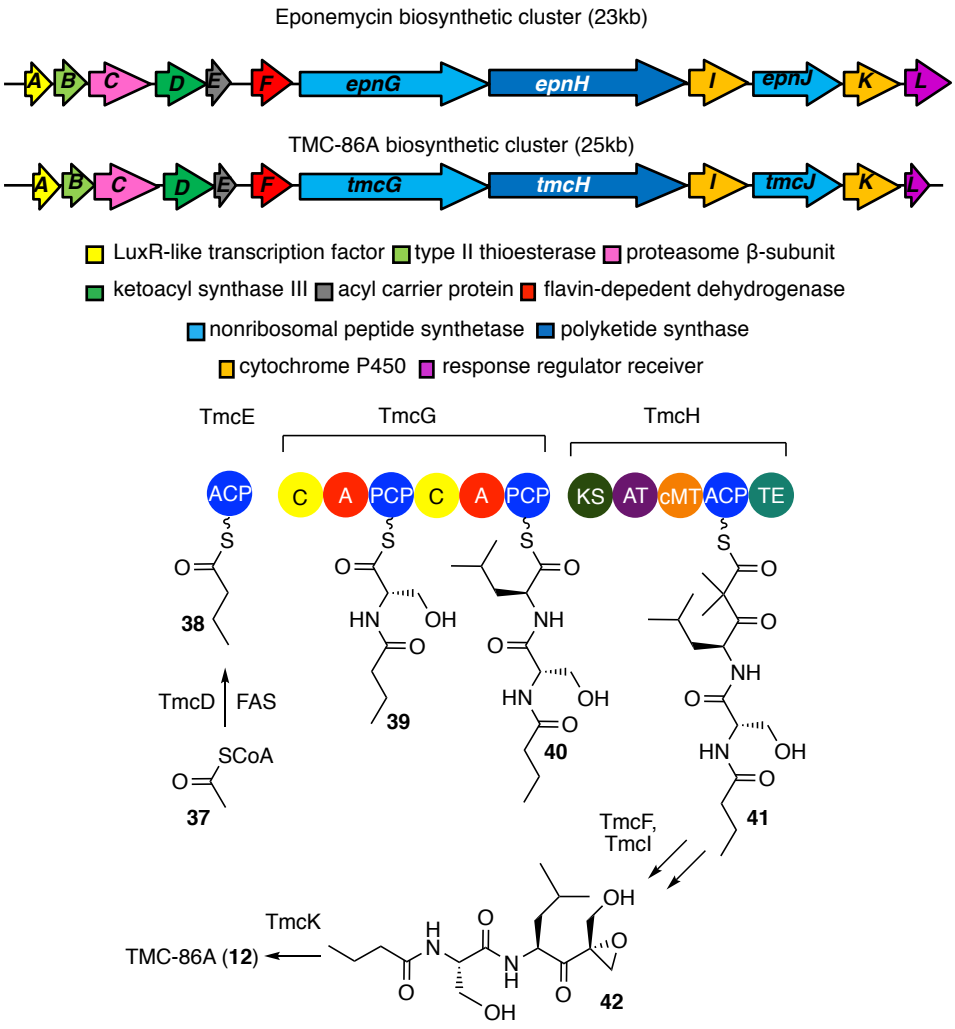


Figure 1.9. A: Arrangement of genes in the eponemycin and TMC-86A biosynthetic gene clusters. B: Proposed biosynthesis of TMC-86A. Exact timing of TmcK desaturation remains unclear.

The PKS module TmcH catalyses transacylation of the peptide intermediate by a malonate extender unit, followed by methylation of the α-position of the β-keto thioester mediated by the C-methyltransferase domain (cMT) with *S*-adenosyl methionine **60** (SAM) as the cofactor.[45] Lastly, serine-residue mediated hydrolysis of

Table 1.2: *In silico* BLAST analysis of genes from the TMC-86A gene cluster, revealing their amino acid length, nearest protein homologue and its accession number on the GenBank database, as well sequence identity/similarity and proposed function.[42]

Gene	Amino Acids	Protein homologue	Accession Number	Identity /similarity	Proposed function
tmcA	70	EpnA, *S. hygroscopicus*	AHB38503	86/98	LuxR-like transcription factor
tmcB	220	EpnB, *S. hygroscopicus*	AHB38504	85/91	Type II Thioesterase
tmcC	285	EpnC, *S. hygroscopicus*	AHB38505	92/95	Proteosome β-subunit
tmcD	340	EpnD, *S. hygroscopicus*	AHB38506	89/94	Ketoacyl synthase III
tmcE	76	EpnE, *S. hygroscopicus*	AHB38507	84/90	Acyl carrier protein
tmcF	557	EpnF, *S. hygroscopicus*	AHB38508	89/92	Flavin-dependent dehydrogenase
tmcG	2167	EpnG, *S. hygroscopicus*	AHB38515	82/88	NRPS
tmcH	2051	EpnH, *S. hygroscopicus*	AHB38509	81/86	PKS
tmcI	431	EpnI, *S. hygroscopicus*	AHB38510	88/92	Cytochrome P450
tmcJ	658	EpnJ, *S. hygroscopicus*	AHB38511	90/93	Adenylation domain
tmcK	408	EpnK, *S. hygroscopicus*	AHB38512	93/96	Cytochrome P450
tmcL	183	T261_3953, *S. lydicus* A02	AJT65613	85/89	Response regulator receiver

the intermediate **41** by the thioesterase (TE) domain of TmcH is proposed to yield the substrate for epoxyketone formation. The initial hypothesis was that epoxidation could be facilitated by either flavin-dependent TmcF, cytochrome P450 TmcI, or a combination of both enzymes.

Interestingly, recent unpublished research by Dr. C. Huang of the Challis group suggests that the *epn* cluster is also a TMC-86A **12** producer, and eponemycin **8** is a shunt metabolite obtained from using the fatty acid 6-methylhepatonoic acid from ordinary *Streptomyces* FAS metabolism. The *epn* cluster was heterologously expressed in *S. albus*, producing approximately equimolar amounts of eponemycin and TMC-86A **12**. Complementation of this transformant with a plasmid that constitutively expressed both *epnD* and *epnE* lead to dramatically increased production of TMC-86A **12**, and only minimal production of eponemycin **8**. The current hypothesis is that the promoter region leading into *epnD* is has mutations which affect production compared to the corresponding region found in the *tmc* gene cluster, ensuring poor expression of *epnD* in *S. hygroscopicus* and therefore significant production of eponemycin **12** using an alternate fatty acid metabolite.

1.3.1.1 Regulation and self-resistance

The TMC-86A **8** biosynthetic gene cluster is regulated by LuxR-like transcription factor, *tmcA*. LuxR transcription factors are known to possess a helix-turn-helix domain in the C-terminal region, which binds to and represses DNA transcription, and an N-terminal region which binds an inducer molecule. In Gram-negative bacteria, this method of transcriptional regulation is moderated by diffusion of acyl-homoserine lactones, however in Gram-positive *Streptomyces* species inter-kingdom communication with many discreet compounds often prompts LuxR induction and subsequent expression of advantageous secondary metabolites.[46]

Encoded by the cluster is an alternate proteasome subunit, TmcC. This subunit has homology to a similar subunit encoded in the biosynthetic gene cluster for proteasome inhibitor Salinosporamide A, produced by *Salinispora tropica*. Amino acid mutations

in the proteasome binding pocket ensure resistance to inhibition. By expression of TmcC, *S. chromofuscus* therefore has self-resistance to the epoxyketone inhibitor it produces.[47]

1.3.1.2 Fatty acid biosynthesis by TmcD

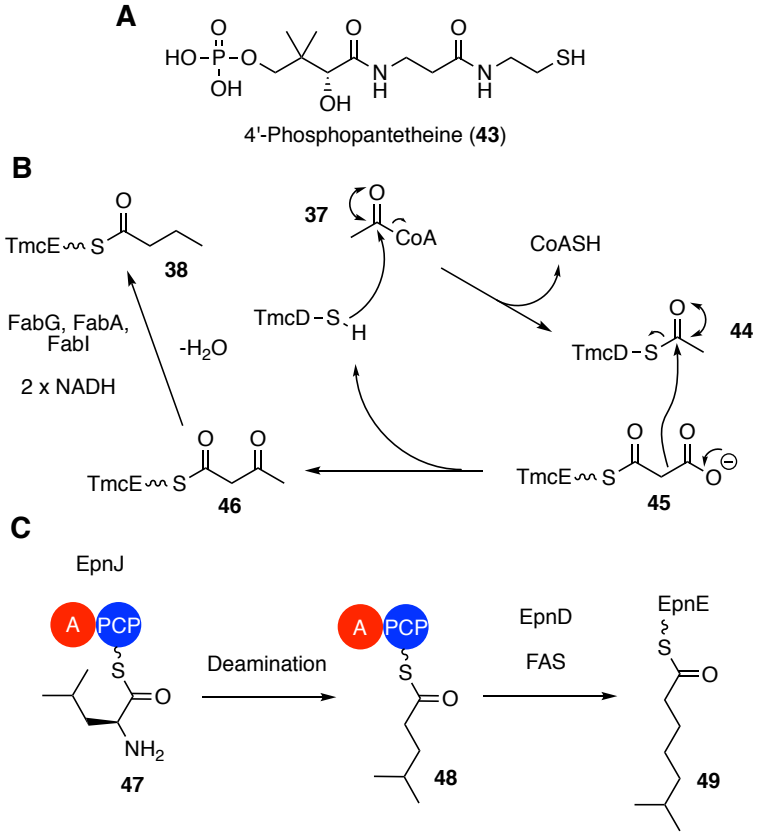


Figure 1.10. A: Structure of phosphopantetheine, PPT. B: Proposed biosynthesis of TMC-86A butyric acid moiety. C: Biosynthesis of 6-methyl heptanoic acid formation proposed by Moore and coworkers.

In order to produce the butyric acid moiety incorporated into the structure of TMC-86A **12**, it is proposed a dedicated FabH-like FAS enzyme TmcD is required. FabH enzymes are responsible for the condensation of conserved cysteine-bound acyl units to a PPT **43** thiol-bound malonyl extender unit (figure 1.10A) of a standalone ACP like TmcE (figure 1.10B).[48] ACP domains facilitate the transacylation of acyl-

intermediates between enzymes. In the case of *Streptomyces* species, their FAS metabolism prefers different acyl-CoA units to the acetyl-CoA **37** required to make butyric acid, often producing branched fatty acids as a consequence.[43] Therefore, a dedicated FAS enzyme with acetyl-CoA **37** specificity is required to produce *β*-keto thioester **46** on TmcE. This is then reduced down to butanoyl-ACP **38** by process of *β*-keto reduction, dehydration and enoyl reduction by FabG, FabA and FabI respectively of existing *Streptomyces* FAS metabolism.[49–51]

In the proposed biosynthesis of eponemycin by Moore *et al*, it is suggested that the incorporated 6-methyl heptanoic alternative fatty acid is a result of deamination of leucine bound to NRPS enzyme EpnJ **47** by an unspecified mechanism, followed by EpnD mediated condensation of malonyl-CoA units to give the 6-methyl heptanoic fatty acid attached to EpnE **49** (figure 1.10C).[41] However, as previously discussed, unpublished data from Dr. C. Huang demonstrate that constitutive promotion of EpnD and EpnE makes the *epn* cluster produce TMC-86A **12** as its major metabolite in *S. albus*. While low levels of eponemycin **8** in this expriment may result from EpnJ not also being constitutively promoted, or possibly from higher levels of butanoyl-CoA being present in *S. albus* compared to the original strain which out-competes 6-methyl heptanoic acid incorporation, it seems unusual for a biosynthetic gene cluster to produce two competing fatty acid precursors. Therefore, it seems more likely 6-methyl heptanoic acid is derived as a branched chain fatty acid from innate *Streptomyces* fatty acid metabolism. An alternate function for both TmcJ and EpnJ under this hypothesis remains unknown, with one possibility being they are non-functional.

1.3.1.3 NRPS domains and peptide synthesis by TmcG

Following butanoyl thioester synthesis, the next stage of TMC-86A assembly is peptidation by the NRPS TmcG. NRPS enzymes are composed of domains serving different catalytic functions organized into modules, making NRPS enzymes multifunctional.[44] The key domains found in NRPS enzymes are:

Adenylation (A) domain – typically around 500 amino acid residues in length, this domain possesses an active site pocket between a small C-terminal subunit and a large

N-terminal subunit.[44] The role of the A domain is to activate an amino acid using ATP to form an aminoacyl adenylate (such as **51**), allowing for reaction with the PPT **43** thiol of a post-translationally modified PCP domain to furnish **52** (figure 1.11A). The A domain is also responsible for amino acid specificity of a module, with variations in eight amino acid residues found in the A domain binding pocket facilitating amino acid recognition.[52]

Peptidyl carrier protein (PCP) domain – approximately 80-110 amino acids in length, the PCP domain is structured as four α-helices with the active site serine located just before the second α-helix. Posttranslational modification of this serine with PPT **43** in turn allows for the domain to facilitate transference of a peptide intermediate (figure 1.11B).

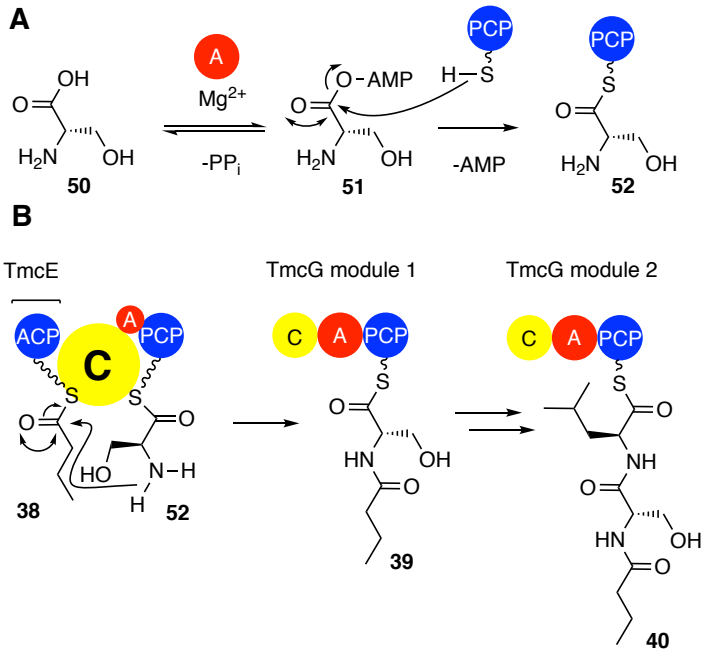


Figure 1.11. A: Adenylation domain catalysed ATP activation of serine, A = Adenylation, PCP = peptidyl carrier protein. B: Condensation (C) domain mediated transpeptidation, resulting in proposed TmcG module 2 intermediate 40.

Condensation (C) domain – usually near 450 amino acids in length, the condensation domain is responsible for catalysing peptide bond formation in an NRPS enzyme between an upstream module and a downstream PCP domain which has been charged with an amino acid (figure 1.11B).[44] C have a nucleophilic docking site which interacts PPT **43** arm of the downstream PCP, and an electrophilic site for the upstream PPT tethered to a PCP, allowing for close association of amino acid and thioester substrates to facilitate formation of the peptidic bond, such as in **38**.

Thioesterase (TE) domain – about 250 amino acids in length the TE domain is responsible for release of a peptide chain from an NRPS module. The active site serine of the TE domain forms an ester linkage with the peptide thioester intermediate bound to the PPT **43** arm of a PCP domain by addition elimination, and then allows for chain release either by hydrolysis or cyclisation using a nucleophilic functionality in the peptide backbone.[44] The catalytic mechanism of the TE domain is similar to that found in PKS enzymes (figure 1.12B).

Tailoring domains – cyclisation (CY) domains are specialised forms of C domains which are responsible for both the condensation of a cysteine, serine or threonine residues and their subsequent cyclisation to form thiazoline or oxazoline, which often have a role in metal chelation when present in natural products.[44] The epimerisation (E) domain of NRPS enzymes is responsible for epimierising the stereocenter of amino acids incorporated into a peptide compound, working in tandem with a C domain which is selective for incorporation of the D-stereochemistry due to E domains producing amino acids as an enatiomeric mixture. Lastly, *C*-methytransferase or *N*-methytransferase (cMT or nMT) domains catalyse methylation of carbon and nitrogen respectively which discourages degradation of peptidic natural products by a protease. This reaction uses the cofactor SAM **60**, and functions similarly to that seen for PKS enzymes (figure 1.12C).[44]

The two modules in TmcG are composed of A, C and PCP domains, and condense L-serine followed L-leucine to fatty acid chain providing **39** and **40** respectively (figure 1.11B).[44]

1.3.1.4 PKS domains and extension by TmcH

The final step of TMC-86A biosynthesis before release as the putative substrate for epoxyketone formation is elongation and dimethylation by the PKS TmcH. Similar to NRPS enzymes, type I PKS enzymes are made of modules formed from domains with differing catalytic activities:

Acyl carrier protein (ACP) domain – similar in size and function to the PCP domains found in NRPS enzymes, with four helical bundles and around 80 amino acids in length, ACP domains carry a catalytic serine residue which is post-translationally modified with PPT. The thiol of the PPT arm is responsible for the shuttling of polyketide intermediates by way of transacylation such as **58** (figure 1.12B).[53]

Acyl transferase (AT) domain – the AT domain is responsible for charging a ACP domain with an extender unit, most commonly malonyl-CoA **53** produced by acetyl-CoA carboxylase. Approximately 300 amino acids in length, in *cis*-AT PKS enzymes the AT domain follows a ketosynthase (KS) domain, whereas in a *trans*-AT PKS there is a standalone AT enzyme. The catalytic serine residue of an AT domain is activated by histidine, which allows for addition-elimination at the malonyl-CoA thioester to give **54**. Following this, nucleophilic addition-elimination of the PPT arm thiol on an ACP domain charges that ACP domain with the extender unit as in **55** (figure 1.12A).[53]

Ketosynthase (KS) domain – responsible catalysis of the carbon-carbon bond forming reaction in PKS enzymes, the KS domain is typically 430 amino acids residues in length and readily forms homodimers with the KS domains of identical PKS enzymes. The key catalytic sites in the domain are a TACSSS motif containing the catalytic cysteine as well as HGTGT and KSNIGHT containing vital histidine residues. The cysteine itself is close to an α-helix, the positive dipole of which assists in increasing the nucleophilicity of the catalytic residue. All this serves to facilitate acquisition of upstream substrate intermediate **56** from a carrier protein by addition-elimination to form **57**, followed by decarboxylative Claisen condensation with a

downstream extender-unit loaded ACP to form a β-keto thioester on that domain and thereby extend the length of the intermediate (figure. 1.11B).[53]

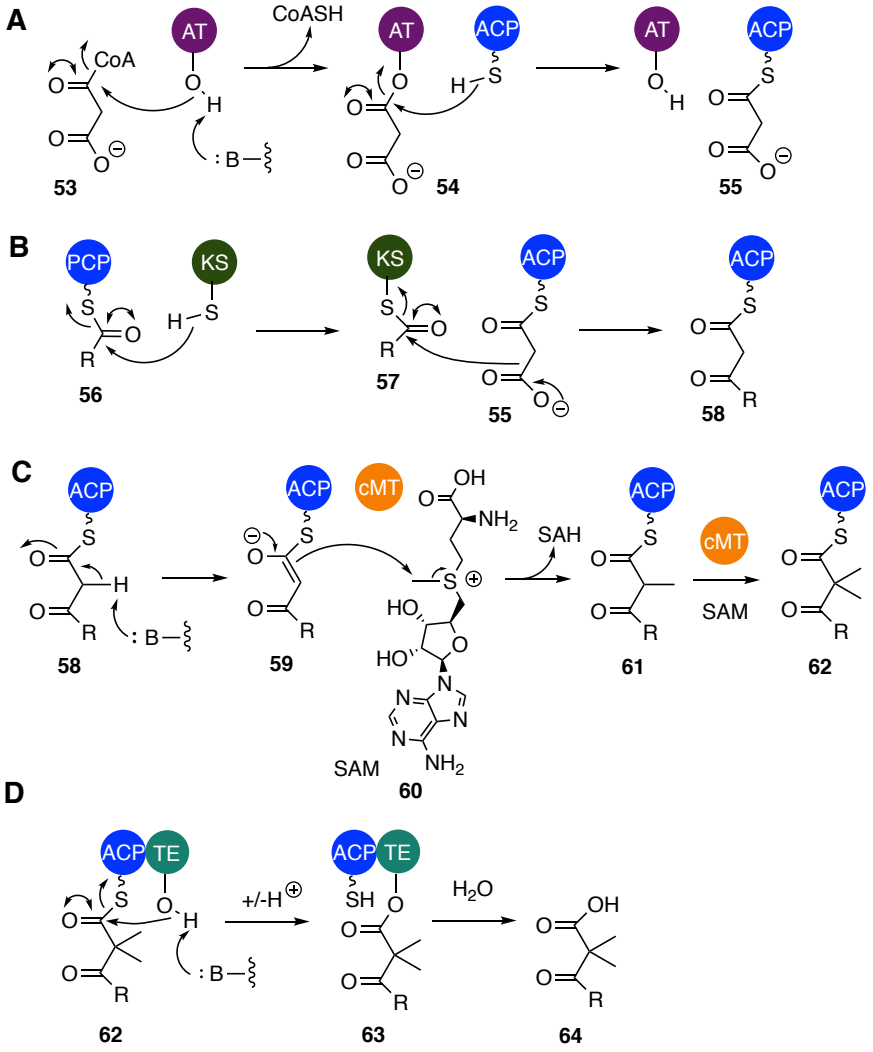


Figure 1.12. A: AT mediated ACP malonylation. B: Transacylation of KS domain, and subsequent decarboxylative Claisen condensation to form an ACP-bound β-keto thioester. C: Dimethylation of β-keto thioester using SAM 60 by cMT domain. D: TE domain hydrolytic chain release of polyketide to give carboxylic acid 40.

***C*-Methyltransferase (cMT) domain** – similar to MT domains found in NRPS enzymes, a cMT transferase methylates the α-position of a β-keto thioester attached

to an ACP domain using SAM cofactor **60** (figure 1.13). cMT domains are comprised of approximately 130 residue N-terminal region which recognises the substrate intermediate, and C-terminal region around 190 residues in length which binds SAM with a characteristic xGxGxG motif. The precise mechanism of deprotonation of intermediates such as **58** to allow for S_N2 attack on SAM by way of enolate **59** is currently unknown (figures 1.11C and 1.12). *O*-methyltransferase (oMT) domains also exist which operate by a similar mechanism to methylate alcohol functionalities.[53]

Thioesterase (TE) domains – similar to TE domains present in NRPS enzymes, this domain is responsible for release of the polyketide intermediate from the PKS enzyme, again either by cyclisation using a nucleophilic residue present in the intermediate or by hydrolysis. Usually ranging from 240-290 amino acids in length, the catalytic serine of a GxSxG motif is base activated and undergoes addition elimination at a thioester **62** attached to the PPT arm of an upstream ACP domain (figure 1.12D). Following this, attack by base activated water or an internal secondary alcohol facilitates release, in the case of TMC-86A resulting in carboxylic acid **64** (figure 1.12D).[45,53]

The PKS module in TmcH is comprised of KS, AT, cMT, ACP and TE domains, and putatively forms carboxylic acid **64** (figure 1.12).

Other catalytic PKS domains – The ketoreductase (KR) domain is typically between 420 and 460 residues in length and is responsible for stereoselective reduction of the β-keto group in a polyketide thioester intermediate using a reduced nicotinamide adenine dinucleotide phosphate (NADPH) cofactor. The presence or absence of an important LDD motif dictates the resulting stereochemistry of the alcohol, with some KR domains also capable of furnishing specific stereochemistry of both the alcohol and α-methyl group in the case of methylation by a cMT domain.[53] With a β-alcohol installed by the KR domain, the dehydratase (DH) domain is then responsible for dehydration to provide an α,β-unsaturated thioester. DH domains are ordinarily around 280 amino acids in length, use catalytic aspartic acid and histidine residues to protonate the alcohol and abstract the α-proton respectively. Lastly, enoylreducatse (ER) domains of around 310 amino acids in length use an NADPH cofactor to reduce

α,β-unsaturated thioester intermediates resulting from action of a DH domain to their saturated counterparts, with the capacity to produce a specific enantiomer in the case of α-substitution.[53]

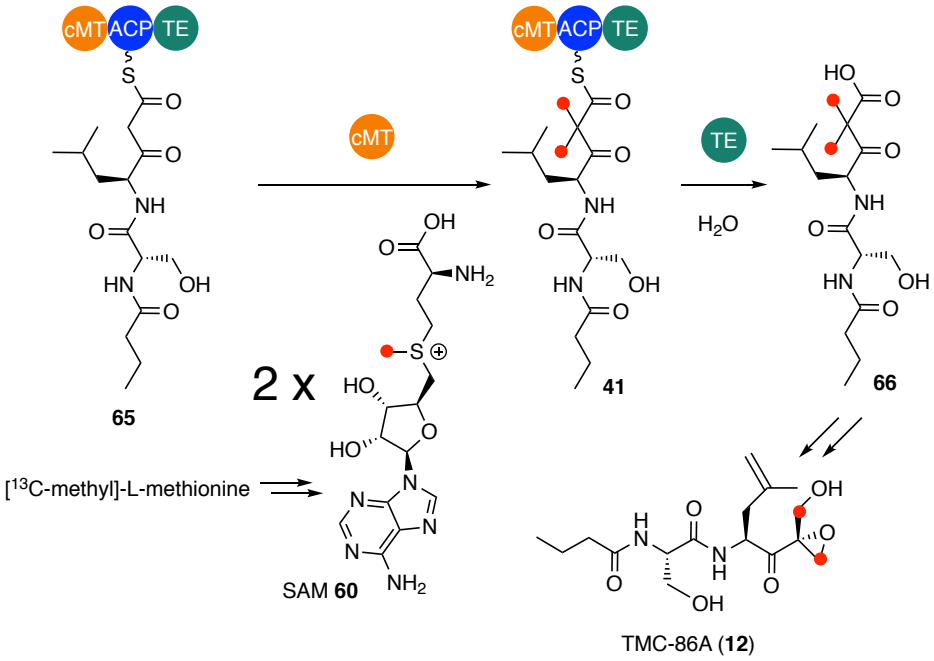


Figure 1.13: Summary of [^{13}C-methyl]-*L*-methionine feeding experiments conducted in *S. chromofuscus* by D. Zabala. Red circles indicate the presence of a [^{13}C]-isotopic label.

To investigate whether the cMT domain was catalysing mono- or di-methylation, Dr. D. Zabala of the Challis group cultured *S. chromofuscus* in the presence of [^{13}C-methyl]-*L*-methionine. Subsequent ^{13}C nuclear magnetic resonance (NMR) spectroscopy revealed incorporation on the isotopic label into both methylene carbons of the α-hydroxymethyl-*α,β*-epoxyketone moiety present in TMC-86A **12**.[42] This indicates that the thioester intermediate bound to the ACP in TmcH is dimethylated by the cMT domain, and that both of these methyl groups remain present while progressing from the putative substrate to TMC-86A **12** (figure 1.13).

1.3.2 Biosynthesis of the TMC-86A pharmacophore

1.3.2.1 Putative substrates

While hydrolytic release of the α-dimethyl-β-keto ester bound to the TE domain of TmcH (figure 1.12D) should theoretically yield α-dimethyl-β-keto carboxylic acid **66**, α-dimethyl-β-keto carboxylic acid compounds are known to be unstable. This is primarily due to the β-keto group providing a natural electron sink and partner for hydrogen bonding with the carboxylic acid, encouraging a favourable conformation.[54] As rate of decomposition had been demonstrated to be independent of the dielectric constant of solvent for these compounds, indicating the presence of a neutral transition state, the most widely agreed upon mechanism of degradation is decarboxylation through a cyclic transition state to produce an enol.[55] Enols **68** can then readily tautomerize to isopropyl ketones **69** (figure 1.14).

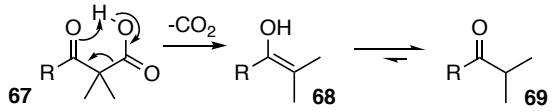


Figure 1.14: Decarboxylative degradation of an α-dimethyl-β-keto carboxylic acid 67 through a cyclic transition state to provide enol 68, and subsequent tautomerization to produce isopropyl ketone 69.

With this in mind, there are two possible products of hydrolysis of the TE domain bound α-dimethyl-β-keto ester **70**: the α-dimethyl-β-keto carboxylic acid **66**, and the isopropyl ketone **71** resulting from its decarboxylation. This decarboxylation could either occur spontaneously, or it may be catalysed by the TE domain itself (figure 1.15). Both putative substrates would need to be synthesised in order to determine the true substrate required for epoxyketone formation.

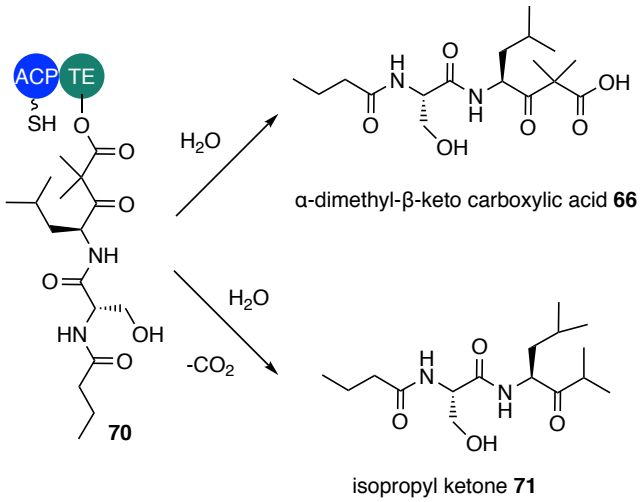


Figure 1.15: The two possible substrates, 66 and 71, resulting from hydrolysis of the α-dimethyl-β-keto ester bound to the TE domain of TmcH.

1.3.2.2 Putative acyl-CoA dehydrogenase TmcF

TmcF has homology to flavin-dependent acyl-CoA dehydrogenase (ACAD) enzymes, with 89% identity and 92% similarity to EpnF. The mechanism of action of ACAD enzymes is concerted fission of two C-H bonds adjacent to fatty acid CoA thioester **73**, with proton abstraction at the α-position facilitated by a conserved glutamate residue and concomitant hydride-equivalent donation from the β-position to a flavin cofactor.[56] In the archetypal case of medium chain acyl-CoA dehydrogenase from FAS, ordinarily the pK_a of the α-position would be too high to allow for deprotonation by glutamate (pK_a > 20), however hydrogen-bonding to the thioester carbonyl by a ribityl-functionality of a cofactor decreases the pK_a of this position to around 8.[56] With a bound substrate, the pK_a of the corresponding glutamic acid is around 9, facilitating the concerted mechanism and using the hydrogen atoms shown, previously demonstrated with deuterium labelling experiments (figure 1.16A). Other natural products which undergo catalysis by enzymes with ACAD homology are pyoluteorin **76** and undecylprodigiosin **77** from *Pseudomonas fluorescens* Pf-5 and *S. coelicolor* A3(2) respectively, where PCP-bound proline is desaturated to a pyrrole moiety.[57,58]

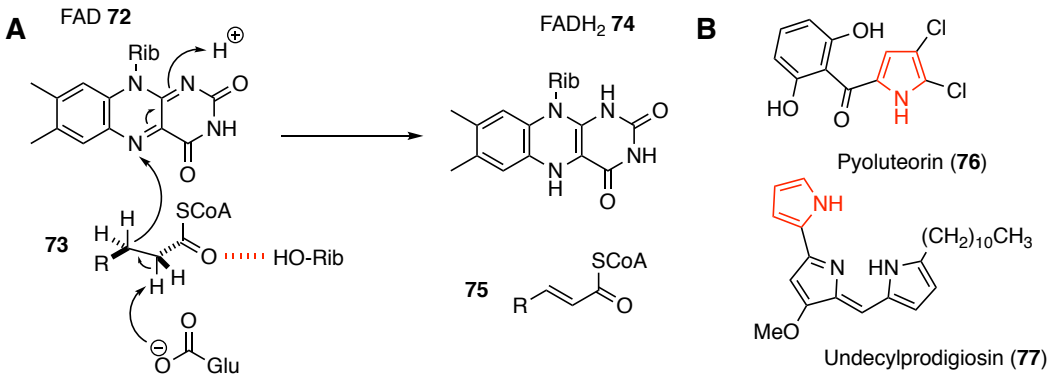


Figure 1.16. A: Mechanism of action of dehydrogenation in medium chain acyl-CoA dehydrogenase, with hydrogen bonding from a ribityl-functionality of a cofactor. B: Structures of pyoluteorin and undecylprodigiosin, natural products dehydrogenated by enzymes with ACAD homology. Moieties highlighted in red demonstrate positions of the catalytic desaturation.

Though medium chain fatty acid-CoA molecules are the typical substrates for ACADs, due to the poor orbital overlap between the sulphur lone pair of electrons and the carbonyl group in a thioester bond, the chemical qualities of a thioester bond are not far removed from that of a ketone such as **71**. One possible route to epoxyketone formation was therefore initial dehydrogenation of **71** by TmcF, using a mechanism directly analogous to the ACAD desaturation of thioesters. This reaction would form the reactive α,β-unsaturated ketone **79**, facilitating epoxidation by the cytochrome P450 TmcI. Alternatively, TmcF was the most likely enzyme to react with the electron-rich carboxylic acid **66** in the instance of **71** not being a viable substrate, through a then unknown mechanism (figure 1.17).

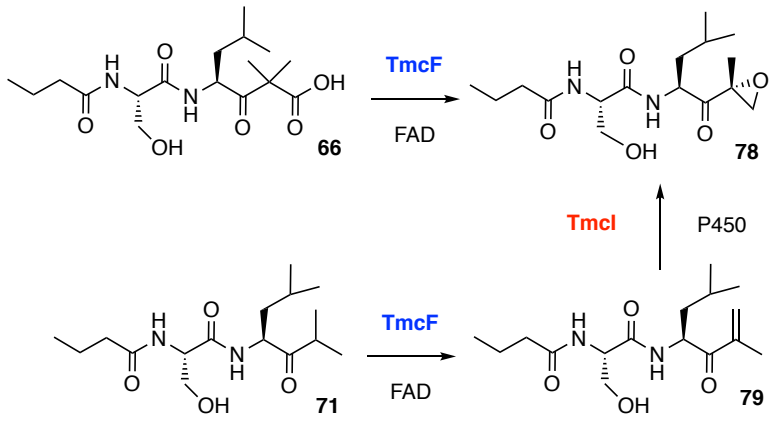


Figure 1.17: The two proposed enzymatic routes to epoxyketone 78.

1.3.2.3 Cytochrome P450 TmcI

With the α,β-unsaturated ketone **79** proposed to form by TmcF dehydrogenation, epoxidation is required to complete the epoxyketone pharmacophore of **78**. Cytochrome P450s are known to catalyse epoxidation reactions with alkene substrates, and of the two P450 enzymes in the cluster it was determined that TmcI was most likely to be responsible for this catalysis, due to it being widely conserved across epoxyketone biosynthetic gene clusters.[59] TmcK by contrast appeared to be only conserved in epoxyketone biosynthetic gene clusters where leucine desaturation occurred, making it putatively responsible for this tailoring reaction.[41]

In order for P450 enzymes to be catalytically functional, the prosthetic heme group bound to a conserved cysteine residue needs to undergo reduction by an electron transport chain as it is incapable of being reduced directly by the necessary cofactors (figure 1.18A).[60] In *Streptomyces* species, electron transport to a P450 is mediated by the proteins ferredoxin (Fdx) and ferredoxin reductase (FDR) from either reduced nicotinamide adenine dinucleotide (NADH) or NADPH (figure 1.18B).[61]

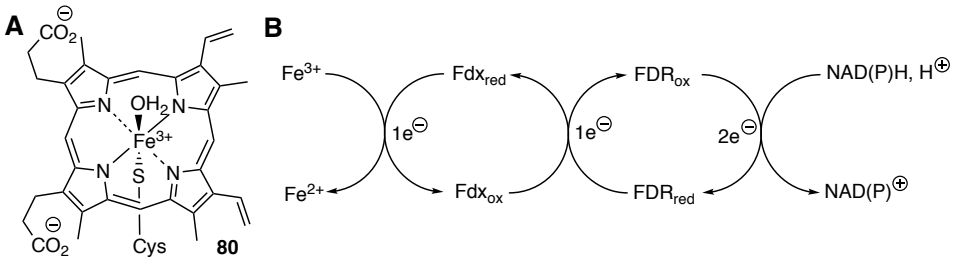


Figure 1.18. A: Structure of low spin ferric resting state of a heme prosthetic group in a cytochrome P450. B: Electron transport chain of a P450, using ferredoxin, ferredoxin reductase and NAD(P)H.

The catalytic cycle of cytochrome P450 enzymes (figure 1.19A) begins with conversion of the six coordinate low-spin ferric resting state of heme to high spin five coordinate ferric resting state upon substrate binding **81**.[59,60] Reduction by an electron transport chain converts iron to its ferrous state **82**, which will reduce oxygen upon binding to form a ferric-superoxo species **83**. Further electron transport chain mediated reduction and subsequent protonation converts this species to ferric peroxide **84**. Additional protonation of this hydroperoxy-species produces water and an oxy-ferryl porphyrin cation radical species **85**, believed to be the responsible for substrate catalysis.[62] Monooxygenation of a substrate by this species, followed by substrate dissociation and the binding of a water molecule, restores the original six coordinate low-spin ferric resting state of heme **80**.[59,60]

Epoxidation by a P450 enzyme from the oxy-ferryl porphyrin cation radical species **85** occurs by two electron oxidation to an alkene bond, though there is ongoing debate as to whether this is by a concerted or stepwise mechanism (figure 1.19B).[59] Epoxide functionalities of several natural products, including epothilone A **89** and mycinamicin II **90** from *Sorangium cellulosum* and *Micromonospora griseorubida* respectively, are putatively produced by this mechanism (figure 1.19C).[63,64]

As previously discussed, the second cytochrome P450 TmcK is putatively responsible for dehydrogenation of the leucine side chain occurring after release from TmcG (section 1.3.1). Following hydroxylation of the leucine side chain by species **85** (figure 1.19D), dehydration would form the carbon-carbon double bond present in TMC-86A.

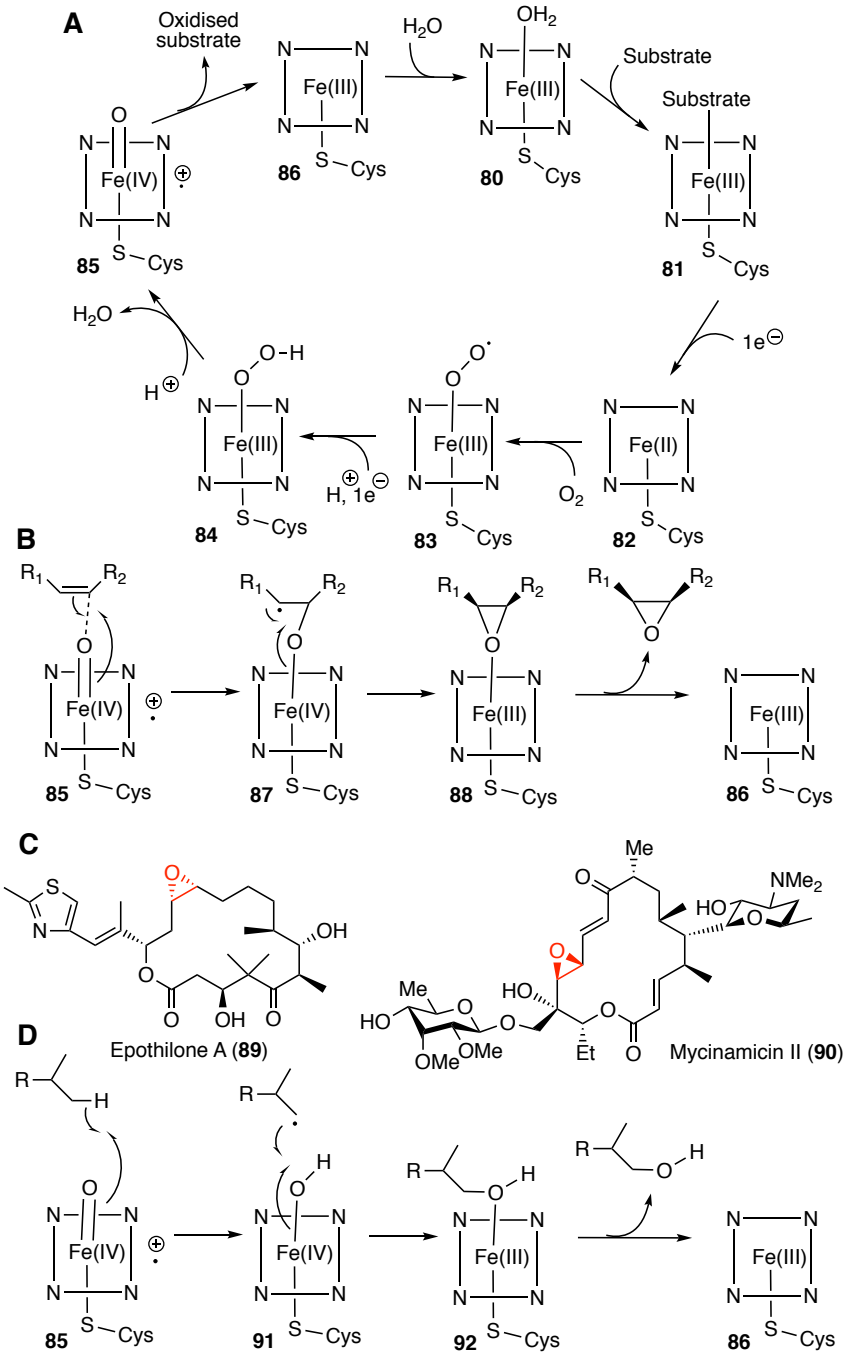


Figure 1.19: A. The catalytic cycle of a cytochrome P450. B: Mechanism of alkene epoxidation by a cytochrome P450. C: Structures of epoxide epothilone A and mycinamicin II. D: Mechanism of hydroxylation by a cytochrome P450, allowing for resulting dehydration.

1.4 Project aims

In the industrial production of epoxyketone proteasome inhibitors, the epoxidation step has proven to have poor stereoselectivity. Synthetic pharmaceuticals were originally designed based on the natural product epoxomicin, though the enzymes responsible for the stereoseletive biosynthesis of the α,β-epoxyketone pharmacophore in these compounds has remained poorly understood. The key aim of this research is therefore to determine which enzymes and substrates accomplish this biosynthetic reaction.

To achieve this, firstly it is necessary to clone, express and subsequently overproduce the cytochrome P450 TmcI and flavin-dependent ACAD-like TmcF. Two putative substrates, the isopropyl ketone **71** and the α-dimethyl-β-keto carboxylic **66** will need to be produced through synthesis. Conducting *in vitro* assays with either **71** or **66** and TmcF or TmcI should reveal the responsible enzyme for α,β-epoxyketone, and therefore produce the TMC-86A analogue **78**.

In the event of successful determination of both substrate and enzyme required for *α,β*-epoxyketone biosynthesis, substrate tolerance of the corresponding enzyme as well as mechanistic interrogation will be conducted.

Chapter 2: Overproduction of recombinant EpnF and TmcI

2.1 Cloning of *tmcF* and *tmcI* into the pET151 vector

To investigate the catalytic activity of the putative flavin-dependent dehydrogenase TmcF and the predicted cytochrome P450 TmcI, the genes encoding these proteins were cloned into an *E. coli* expression vector. Utilizing genomic DNA isolated from *S. chromofuscus* ATCC49982 and appropriate primers, polymerase chain reactions (PCRs) were conducted on the DNA to produce amplimers corresponding to each gene. The PCR products were visualised using agarose gel electrophoresis, and fragments of the correct size (1.3 kb and 1.7 kb for TmcI and TmcF respectively) were excised from the gel and purified (figure 2.1).

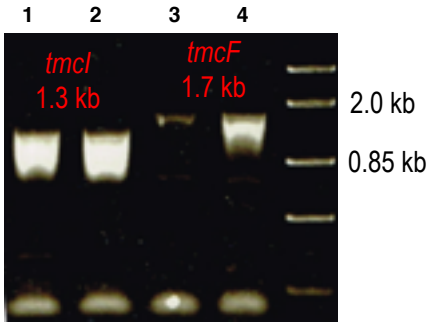


Figure 2.1: Agarose gel of *tmcI* (lanes 1 and 2) and *tmcF* (lanes 3 and 4) amplimers separated by electrophoresis. PCR was conducted at 55 °C in the first lane and 60 °C in the second lane for each amplimer, with a 1 kb marker in lane 5.

Amplimers were cloned into pET151, which has several features that facilitate the overproduction and purification of a recombinant proteins encoded by the gene inserted into it (figure 2.2). The *lacI* gene of the vector encodes the *lac* repressor, which binds to the *lac* operon in the cloning site and controls transcription. This feature allows for induction by isopropyl β-D-1-thiogalactopyranoside (IPTG), a stable mimic of the natural inducer allolactose, which releases the *lac* repressor from the operon.[65] The T7 promoter binds the highly specific T7 RNA polymerase encoded by DE3 *Escherichia coli* cell lines, enabling very efficient transcription of the gene of interest upon induction.[66] The introduction of a start codon upstream of the cloning site allows for expression of recombinant proteins with a hexahistidine (His_6) tag

fused to their N-termini facilitating purification by immobilized metal affinity chromatography (IMAC). The ribosome binding site upstream of the start codon ensures efficient translation of the resultant mRNA. Lastly the *bla* gene functions as a selectable marker for ampicillin, ensuring that only cells containing the vector (with or without an insert) will grow in culture media containing ampicillin.

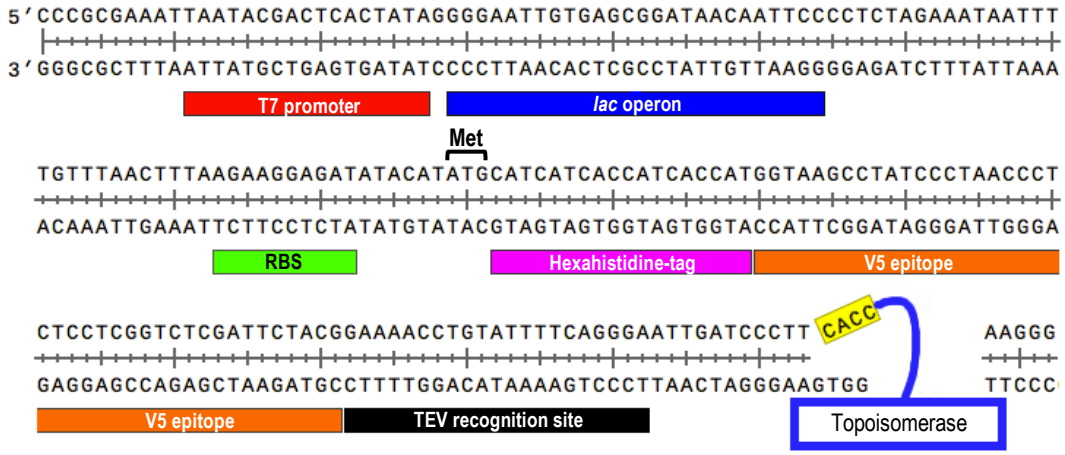


Figure 2.2: Features of the pET151 vector (Invitrogen) in the vicinity of the cloning site. RBS denotes the ribosome binding site, and Met signifies the methionine start codon.

The resulting constructs were used to transform chemically competent *E. coli* TOP10 cells, and positive colonies were selected for by plating on Luria-Bertani (LB) agar containing ampicillin. Single colonies were picked and grown in LB liquid medium containing ampicillin, and the resulting plasmids were isolated from pelleted cells (see experimental section for details).

The integrity of the isolated plasmids was verified using restriction digests, with restriction enzymes selected that would cut once inside the gene of interest and once in the vector backbone (figure 2.3). This provided a distinctive pattern of DNA bands that confirmed the correct orientation of the gene in the vector, approximately corresponds to the kb markings in the kb marker. Plasmids yielding the anticipated digestion pattern were sent for DNA sequencing, and one correct clone was selected in each case for protein overproduction.

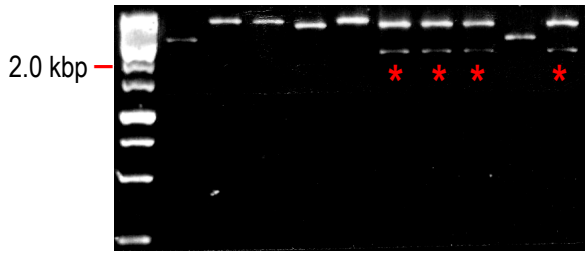


Figure 2.3: Agarose gel electrophoretic analysis of plasmids isolated from ampicillin-resistant clones after digestion with *Pvu*I. Clones containing *tmcI* (lanes marked with an asterisk) are expected to yield bands of 2.2 and 4.8 kb. A kb marker is in the left most lane.

2.2 Overproduction of recombinant proteins in *E. coli* and subsequent purification

2.2.1 Overproduction of TmcF and TmcI

To overproduce TmcF and TmcI, the plasmids containing the corresponding genes were used to transform *E. coli* BL21 (DE3). Transformants were selected by growth on ampicillin LB agar plates, and single colonies were picked and used to produce mini-cultures by overnight growth in liquid LB medium. The mini-cultures were used to inoculate large scale cultures, which were grown to an optical density at 600 nm (A_{600}) of 0.5-0.7, during which period the cells are in the exponential growth phase. Expression was induced by addition of IPTG, allowing for efficient translation and transcription as discussed previously (section 2.1). To facilitate incorporation of the heme prosthetic group into TmcI, 5-aminolevulinic acid was added to the cultures at the same time as IPTG. 5-Aminolevulinic acid is a precursor of the porphyrin ligand incorporated into heme and has been reported to boost the levels of heme incorporation into cytochrome P450 enzymes.[67]

Once induced, the cells were incubated overnight at 15 °C then pelleted by centrifugation. The cell pellet was re-suspended in buffer solution and the cells were lysed (see experimental procedures for details). The insoluble cell debris was separated from the soluble protein fraction by centrifugation.

The recombinant His_6 fusion proteins were purified by IMAC on a nickel-nitrilo acetic acid column. The His_6 tag binds specifically to the immobilised Ni^{2+}, allowing other proteins in the lysate to flow through. Successive washing with buffer containing low and high concentrations of imidazole elutes non-specically bound proteins followed by the recombinant His_6 fusion protein (section 7.2.6).

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) enables visualisation and screening of protein purity at different stages of the purification process. This allows problems such as poor solubility or a lack of protein binding to the nickel column due to burial of the His_6 tag to be identified. SDS both denatures and binds to proteins giving them a uniform *m/z* ratio, allowing for separation on the basis of size. Coomassie Brilliant Blue dye was used to visualise proteins on the gel, which stabilises the blue anionic form of the dye. While the His_6-TmcI fusion was found to be soluble, His_6-TmcF was localised primarily in insoluble inclusion bodies (figure 2.4).

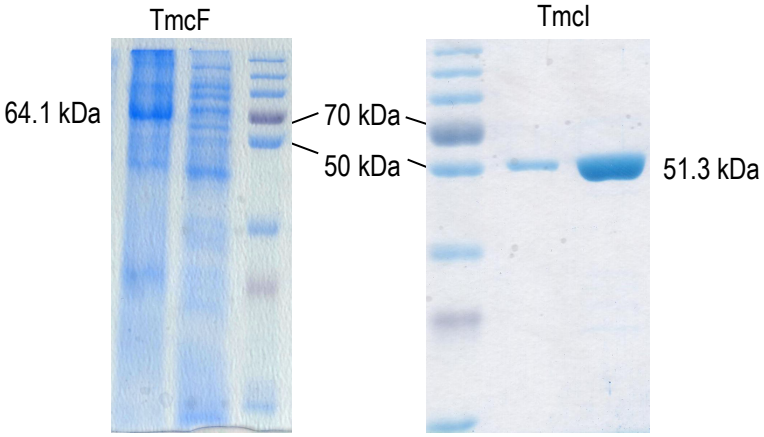


Figure 2.4: SDS-PAGE analysis of insoluble (left lane) and soluble (middle lane) fractions resulting from overproduction of His_6-TmcF (64.1 kDa) in *E. coli* (left), and two different concentrations of purified recombinant His_6-TmcI (51.3 kDa) (right). Molecular weight markers are in the right and left lanes, respectively.

To overcome the problem with the insolubility of His_6-TmcF, expression in *E. coli* Rosetta (DE3) was attempted. This strain contains genes encoding tRNAs for rare

codons and is useful for expression of genes with high G + C content, such as those derived from *Streptomyces* species.[68] The *tmcF* gene was also cloned into the pColdI vector, which facilitates induction of expression at 4 °C by action of the *cspA* promoter, though action of *cspA* is poorly understood.[69] The use of low temperatures during induction of gene expression is known to increase both the solubility and production level of the corresponding protein. Neither approach resulted in the production of soluble His_6-TmcF. Thus we sought to overproduce homologues encoded by other epoxyketone biosynthetic gene clusters.

2.2.2 Overproduction of EpnF

Both the structures and the biosynthetic gene clusters of TMC-86A **12** and eponemycin **8** (produced by *S. hygroscopiscus* ATCC53709) are very similar, and the eponemycin biosynthetic gene cluster encodes a homologue of TmcF (EpnF) with 89% identity.[41,42] Indeed, recent research by the Challis group has shown that heterologous expression of the eponemycin biosynthetic gene cluster in *S. albus* results in the production of approximately equimolar quantities of eponemycin **8** and TMC-86A **12** (sections 1.3.1 and 1.3.1.2). We thus attempted to overproduce EpnF in

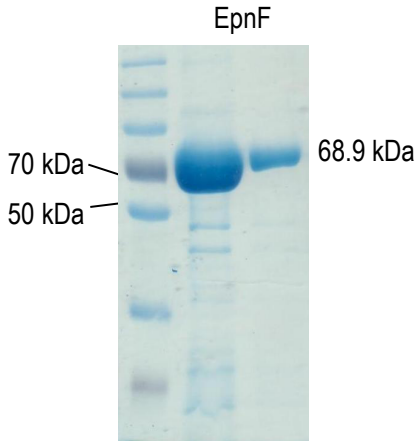


Figure 2.5: SDS-PAGE analysis of two different concentrations of purified His_6-EpnF (68.9 kDa) (middle and right lanes). Molecular weight markers are in the left lane.

E. coli and following the procedures used for TmcF and TmcI (sections 2.1 and 2.2.1) resulted in the production of soluble His_6-EpnF, which was purified using IMAC to near-homogeneity (figure 2.5).[42]

To optimise the overproduction of EpnF, different concentrations of IPTG were evaluated (25/50/100/250 μM). It was determined that 100 μM OF IPTG produced the greatest quantity of EpnF per litre of cell culture. Riboflavin was also added at the same time as IPTG, as this has been reported in the literature to promote the incorporation of the flavin cofactor into flavoproteins.[70]

2.2.3 Characterisation of recombinant proteins

To confirm the identity of the purified recombinant proteins they were analysed by electrospray ionisation-quadrupole-time of flight mass spectrometry (ESI-Q-TOF-MS; by Dr. Matthew Jenner). The masses observed in the deconvoluted spectra obtained from these experiments were consistent with those calculated for His_6-EpnF and His_6-TmcI (figure 2.6), each in the absence of their cofactor. In addition, peptide mass fingerprinting was conducted on each of the purified proteins. Protein bands were excised from an SDS-PAGE gel and digested with trypsin. MALDI-TOF gave 15 and 30 fragments for TmcI and EpnF respectively, corresponding to 33% and 37% sequence coverage (see appendix).[42]

Gel filtration chromatography was conducted on each protein to determine their oligomerisation states. By comparison to the elution volume of protein standards of known molecular weight (section 7.2.9), which have a linear relationship, it was determined that the elution volume of TmcI corresponded to a monomer whereas that of EpnF corresponded to a dimer (figure 2.7).[42] This is consistent with cytochromes P450 being widely observed as monomeric species.[71]

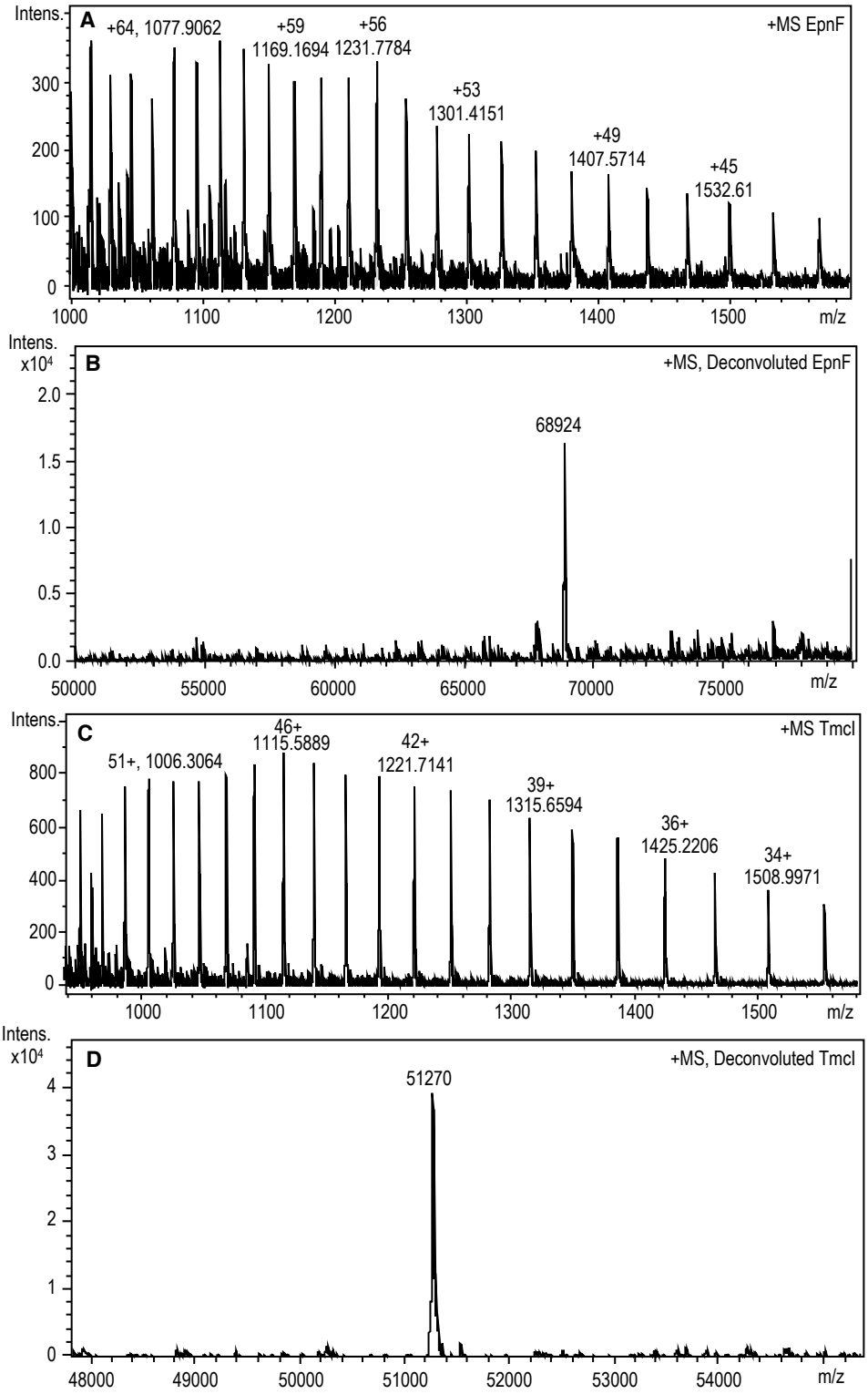


Figure 2.6. Measured (A, C) and deconvoluted (B, D) ESI-Q-TOF mass spectra for His_6-EpnF (*m/z*+ calculated: 68921, observed: 68924) and His_6-TmcI (*m/z*+ calculated: 51269, observed 51270), respectively.

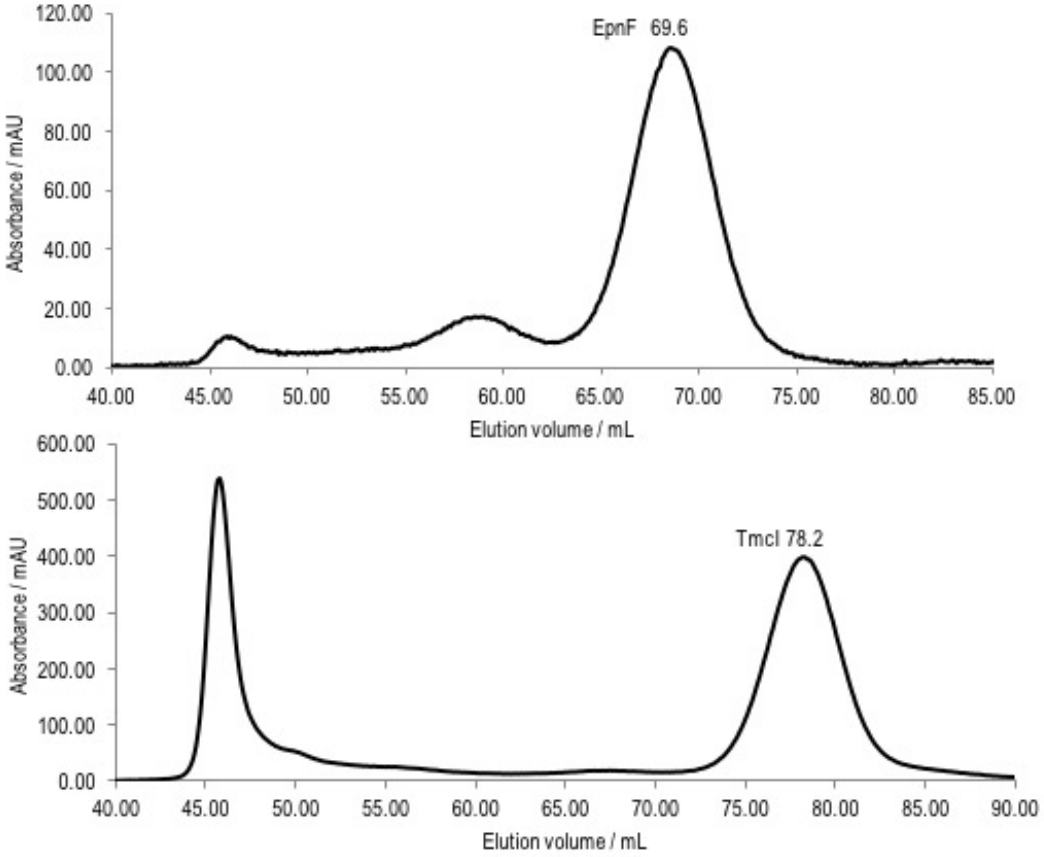


Figure 2.7: Chromatograms from gel filtration analysis of His_6-EpnF (top) and His_6-TmcI (bottom), demonstrating elution volumes of purified proteins corresponding to absorbance maxima.

2.2.3.1 Characterisation of TmcI heme prosthetic group

The observation of a species with *m/z* = 616.1765 (corresponding to the positively charged molecular ion for heme) in the ESI-Q-TOF mass spectrum of His_6-TmcI (figure 2.8) indicated that it contains a heme prosthetic group, as expected for a cytochrome P450. The purity index of His_6-TmcI was determined by first removing imidazole from the protein, and then measuring its UV-visible absorbance spectrum between 250 and 500 nm. The absorbance at 416 nm corresponds to the Soret band of the hexacoordinate low-spin iron(III) in the heme cofactor. A value between 1.0 and 1.9 for the ratio of A_{280}/A_{416} is expected for cytochrome P450 enzymes. The range of

values observed results from differences in the extinction coefficient at 280 nm for individual proteins. Impurities such as imidazole and other proteins will skew this ratio to higher values.[42,72] In this instance, a value of approximately 1.0 was obtained (figure 2.8B).

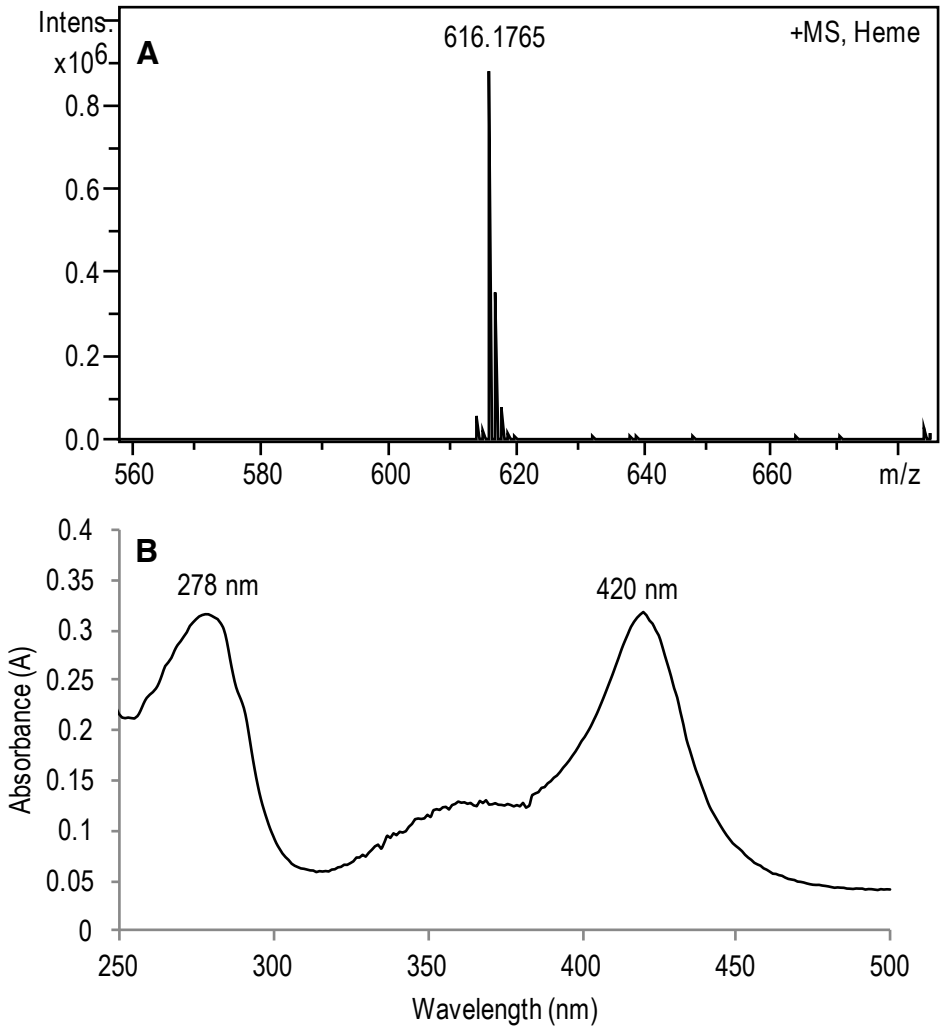


Figure 2.8. A: Mass spectrum of heme prosthetic group released from His_6-TmcI ESI-Q-TOF MS analysis (*m/z* calculated for $C_{34}H_{32}FeN_4O_4^+$: 616.1767, observed: 616.1765). B: UV-visible absorbance spectrum of His_6-TmcI (500 μM) between 250 and 500 nm.

Formation of the characteristic high-spin carbon monoxide complex, absorbing at 450 nm, was attempted using sodium dithionite as a reductant and tricarbonylchloro(glycinato)ruthenium II (carbon monoxide releasing molecule 3 – CORM3), which releases carbon monoxide in aqueous solution. However, this proved problematic, probably due to the presence of oxygen in the buffer solutions. The mass spectrometric and UV-vis spectroscopic analyses of TmcI were judged to have provided sufficient evidence for the presence of a heme prosthetic group in the protein for the purposes of proceeding with investigation into its catalytic function.

2.2.3.2 Characterisation of EpnF flavin cofactor

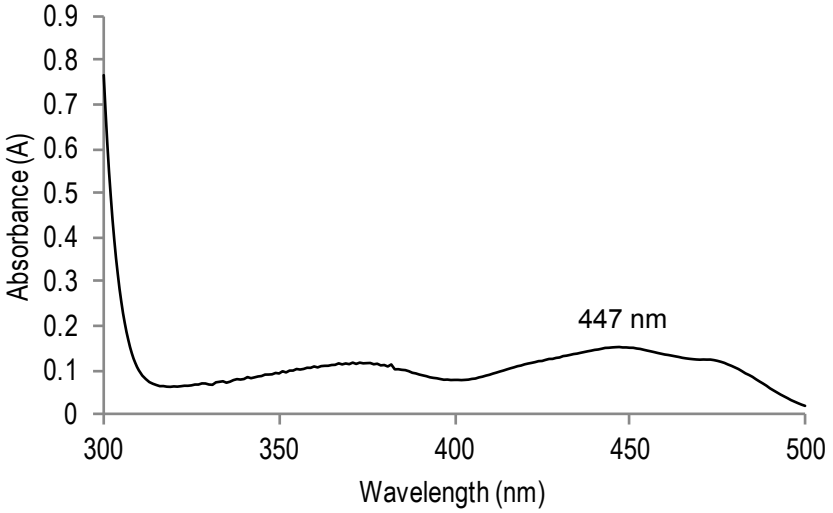


Figure 2.9: UV-visible absorbance spectrum of His_6-EpnF (500 μM) between 300 and 500 nm.

The observed mass of His_6-EpnF in ESI-Q-TOF spectra showed that the protein did not contain any covalent modifications. The flavin cofactor was therefore judged to be non-covalently bound to the active site of the protein. To illuminate the nature of the flavin cofactor, the absorbance spectrum of His_6-EpnF was measured between 300 and 500 nm. An absorbance maximum at 447 nm was observed (figure 2.9).[42] The absorbance maximum of flavin adenine dinucleotide (FAD) is at 450 nm, whereas that for flavin mononucleotide (FMN) is 446 nm. The spectroscopic analysis of EpnF thus

failed to provide conclusive insight into the nature of the flavin cofactor.[73] Heat-denaturation of EpnF, followed by liquid chromatography-mass spectrometry (LC-MS) analysis of the supernatant was attempted. However no ions corresponding to either cofactor were observed, presumably because the quantities of cofactor released were below the limit of detection.

In later research, FAD and FMN were added separately to *in vitro* assays with EpnF. An increase in the quantity of product formed relative to a control reaction without additional flavin cofactor was only observed in the case of FAD (section 4.2.7.3). These data strongly suggest that FAD is the flavin cofactor used by EpnF.

Chapter 3: Synthesis of potential substrates for EpnF and *in vitro* assays

3.1 Synthesis of potential substrates

3.1.1 Choice of substrate analogues

While TMC-86A produced by *S. chromofuscus* has a peptidic *N*-butanoyl-L-seryl-L-4,5-dehydroleucine motif appended to the epoxyketone moiety, the timing of desaturation of the leucine sidechain with respect to pharmacophore installation remains unclear. Therefore, α-dimethyl-β-keto carboxylic acid **66** and the isopropyl ketone **71** were first assessed as potential leucine-derived substrate analogues (figure 3.1). Both substrates could be produced following TE domain mediated hydrolysis from the modular PKS multienzyme TmcH, providing **66** directly or **71** after decarboxylation (section 1.3.2.1). This decarboxylation could either occur spontaneously or through catalysis by the TmcH TE domain.

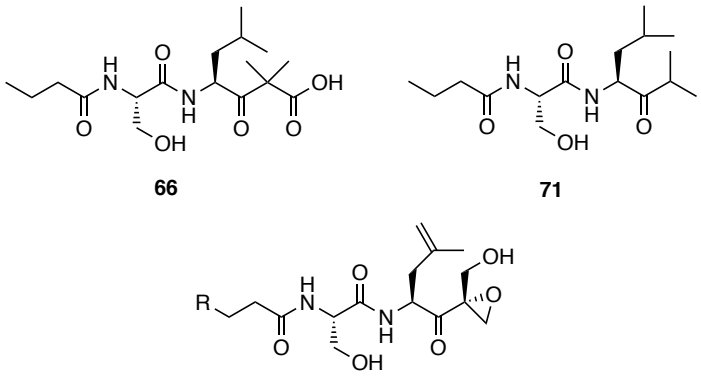


Figure 3.1: Structures of potential substrates isopropyl ketone 66 and α-dimethyl-β-keto carboxylic acid 71, and the structurally-related natural products TMC-86A and eponemycin.

The synthetic work detailed in this chapter was conducted concurrently with enzyme purification, and as detailed previously only EpnF was overproduced in soluble form, with TmcF residing in insoluble inclusion bodies (sections 2.2.1 and 2.2.2). At the time of conducting the research it was not known that the eponemycin biosynthetic gene cluster could also direct the production of TMC-86A when expressed in *S. albus*

(section 1.3.1). However, as eponemycin only differs from TMC-86A by incorporation of a methyl 6-heptanoyl group in place of the butanoyl group (figure 3.1), **66** and **71** were expected to be viable substrates for EpnF.

3.1.2 Synthesis of putative isopropyl ketone substrate 71

3.1.2.1. Proposed synthetic route

To develop a synthetic route to **71**, the compound was retrosynthetically divided into amine **94** and carboxylic acid **93**, and routes to these peptide coupling partners then devised on the basis of existing epoxyketone total syntheses reported in the literature (figure 3.2).

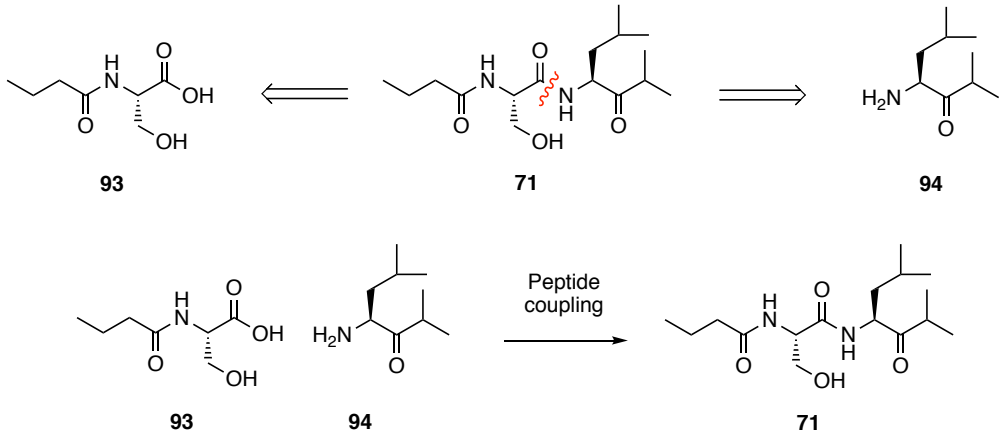
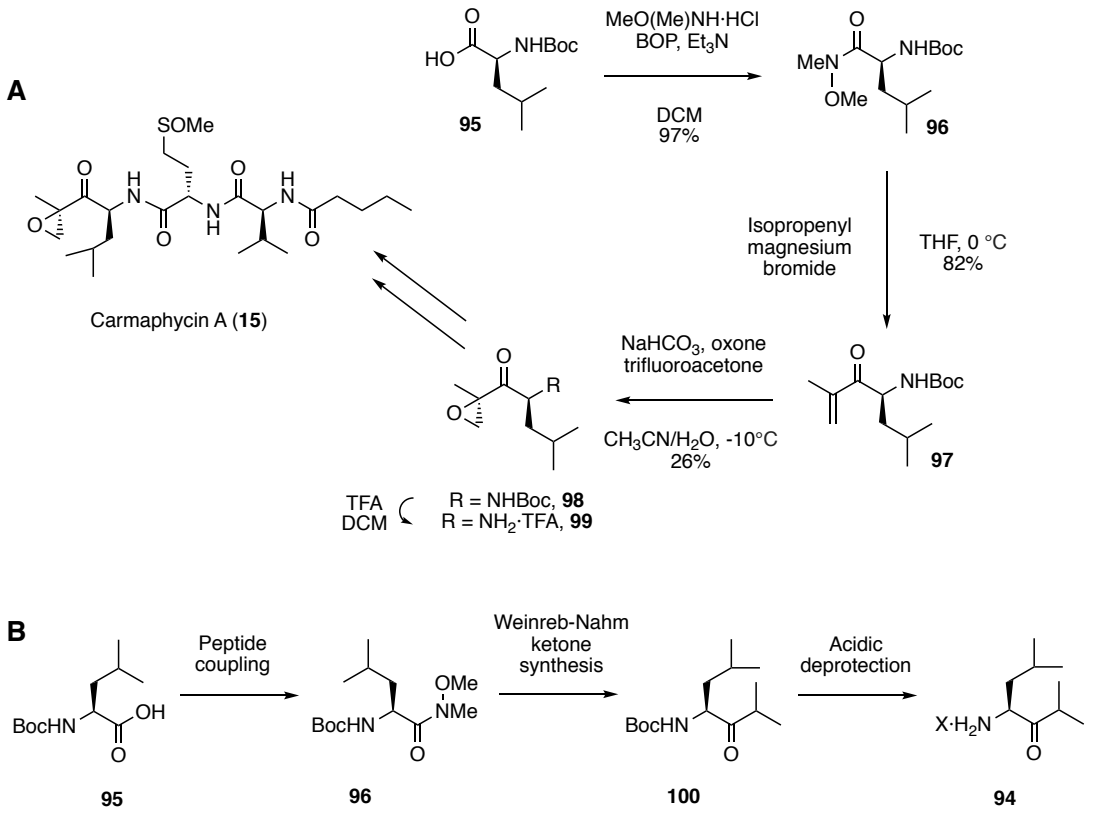


Figure 3.2: Disconnection of the proposed isopropyl ketone substrate analogue for EpnF gives carboxylic acid 93 and amine 94 coupling partners to produced 71.

The natural product carmaphycin A **15** has a leucine residue in an analogous position to **71**, so as a starting point for synthesis of **94** the total synthesis of **15** was reviewed.[22] To synthesise the epoxyketone peptide coupling partner **99**, Weinreb-Nahm ketone synthesis was used to produce α,β-unsaturated isopropenyl ketone **97** via Weinreb amide **96** (for further discussion, see section 3.1.2.3).[74] The epoxyketone **98** was synthesised through base-catalysed epoxidation of **97**, and finally *tert*-butyloxycarbonyl (Boc) deprotection by TFA was used to liberate amine **99** (scheme 3.1A). The total synthesis carmaphycin A **15** was completed by coupling **99** to the

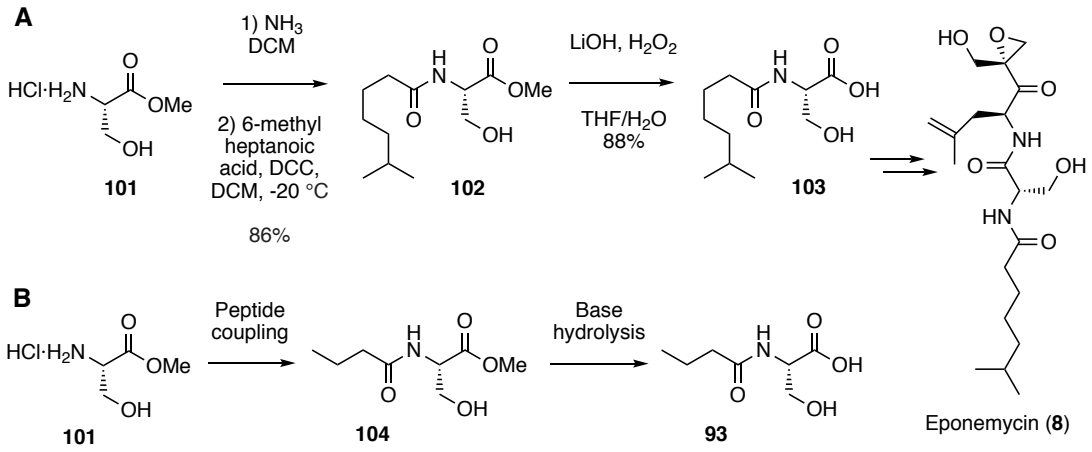
corresponding carboxylic acid. To synthesise **94**, an isopropyl Grignard reagent could be used during Weinreb-Nahm ketone synthesis to install isopropyl ketone functionality and give **100**, forgoing epoxidation. Acidic Boc deprotection would then yield the desired peptide coupling partner **94** (scheme 3.1B).



Scheme 3.1. A: Total synthesis of carmaphycin A, with a single diastereomer purified by HPLC.[22] **B: Proposed synthetic route to amine coupling partner 94.**

As previously mentioned, eponemycin contains an *N*-acyl-L-seryl moiety (figure 3.1). Given this similarity to desired carboxylic acid **93**, the total synthesis of eponemycin was evaluated.[75] Derivatized serine **103** was produced by *N,N′*-dicyclohexylcarbodiimide (DCC)-mediated coupling of L-serine methyl ester **101** to 6-methyl heptanoic acid to produce **102**, followed by basic hydrolysis of the methyl ester with $LiOH/H_2O_2$ (scheme 3.2A). Subsequent steps, including peptide coupling

and epoxidation, give eponemycin. Exchanging 6-methyl heptanoic acid for butyric acid should allow for synthesis of **93** by the same method (scheme 3.2B).



Scheme 3.2. A: Key synthetic steps in eponemycin total synthesis.[75] **B: Proposed synthetic route to 93.**

Peptide coupling reactions building on the *C*-terminus, as proposed between **94** and **93** (figure 3.2), are typically avoided due to a degree of epimerisation from oxazolone **106** formation (figure 3.3). However, epimerisation of the leucine stereocenter was already unavoidable, due to keto-enol tautomerization in the proposed acidic Boc deprotection as well as during the basic conditions required for peptide coupling (scheme 3.1).[76] Taking this into account, a convergent synthesis was retained in the hopes of increasing overall yield relative to a linear synthesis by reducing total number of steps.

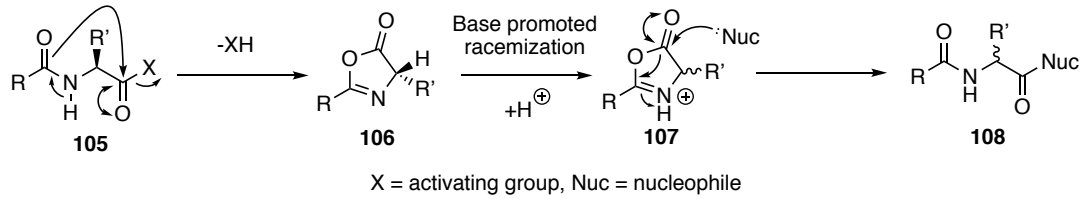
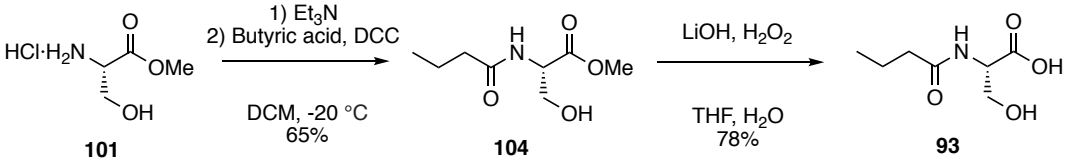


Figure 3.3: Mechanism of oxazolone mediated epimerization during peptide coupling.

3.1.2.2. Synthesis of *N*-butanoyl-L-serine 93

To couple butyric acid to L-serine methyl ester hydrochloride, a DCC-mediated coupling procedure was adapted from the total synthesis of eponemycin (scheme 3.2).[75] While the total synthesis of eponemycin used ammonia gas as a base to convert the amine salt and carboxylic acid to the amine and carboxylate necessary for peptide coupling, for the sake of practicality and safety, triethylamine was chosen for the synthesis of **104**. The initial yield of **104** was low (12%). However, repeating the reaction under an atmosphere of argon and with freshly distilled triethylamine and butyric acid allowed for an improvement to 65%. The coupling was also repeated with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) hydrochloride **110** with hydroxybenzotriazole (HOBt) **115** as an additive. However, a yield of only 62% was achieved. The synthesis of **93** was completed using lithium hydroxide, with the addition of a catalytic amount of hydrogen peroxide (scheme 3.3).[75]



Scheme 3.3: Synthesis of *N*-butanoyl-L-serine 93 through DCC coupling and subsequent ester hydrolysis.

The role of coupling reagents such as the carbodiimides DCC **109** and EDC **110** in peptide synthesis is to activate the carboxylic acid towards addition-elimination by an amine (figure 3.4A).[76] Without a coupling reagent, an unreactive ammonium carboxylate salt forms instead of the desired peptide bond. For carbodiimides, the thermodynamic driving force behind the coupling reaction is the formation of a stable urea. Additives such as HOBt **115** or 4-dimethylaminopyridine (DMAP) **116** can be added to ensure minimization of *N*-acylurea byproducts **114** (figure 3.4B). The rate at which these additives react is faster than rearrangement of the initially formed carbodiimide adduct of the carboxylic acid to form an *N*-acylurea **114**, while still providing an adequate leaving group to ensure successful acyl transfer to an amine.[76]

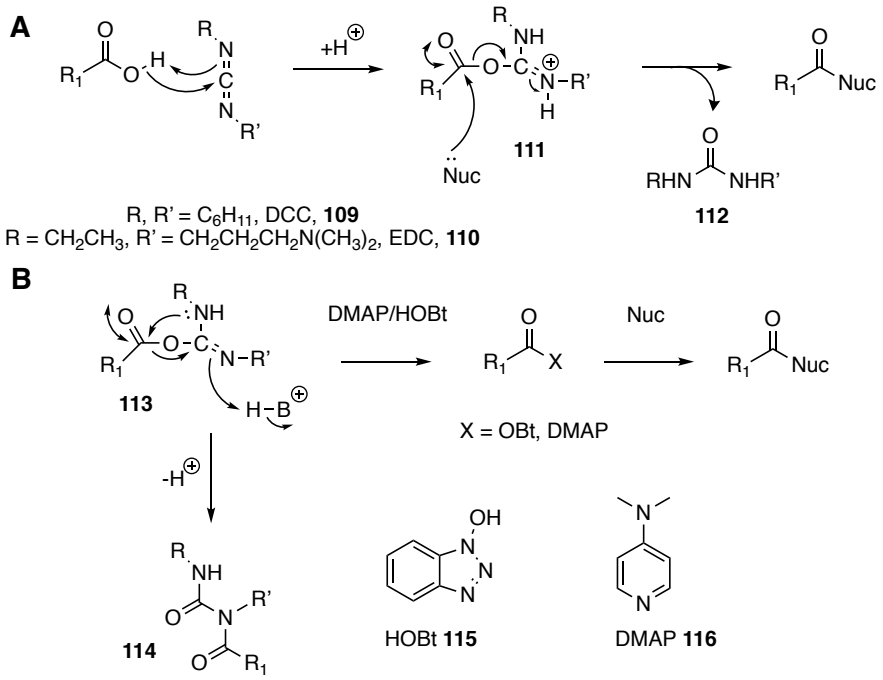
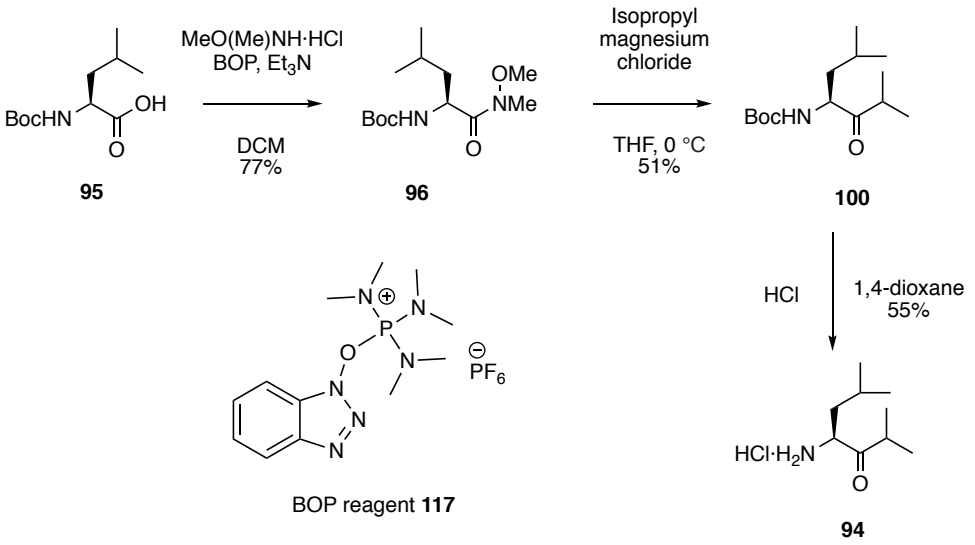


Figure 3.4. A: Proposed mechanism of carbodiimide-mediated peptide coupling. B: Structure of HOBt 115 and DMAP 116, additives used in peptide coupling reactions to reduce the formation of *N*-acylureas 114 via rearrangement of the initially-formed carbodiimide-carboxylate adduct 113.[76]

3.1.2.3. Synthesis of 94

Following the procedure used in the total synthesis of carmaphycin A **15**, coupling of *N,O*-dimethylhydroxyamine to *N*-BOC-L-leucine **95** using (benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate (BOP) **117** was attempted, producing a yield of 77% of the Weinreb amide **96** (scheme 3.4). BOP **117**, similar to the carbodiimide reagents employed previously (section 3.1.2.2), assists peptide coupling by activation of the carboxyl group to addition-elimination by a nucleophile, with formation of a phosphorous-oxygen double bond providing a thermodynamic driving force.[76]



Scheme 3.4: Synthesis of 94 through Weinreb-Nahm ketone synthesis and subsequent Boc deprotection.

Weinreb amides such as **96** are used to ensure only mono-addition of the Grignard reagent to the carbonyl functionality by formation of a stable metal chelate intermediate **119**, which is converted to the corresponding ketone via a low temperature acidic workup (figure 3.5).[74] Using this method with isopropylmagnesium chloride resulted in a yield of 51% of **100**, which could likely be

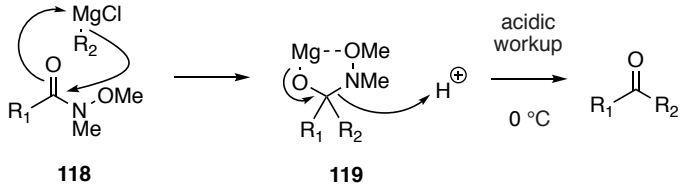


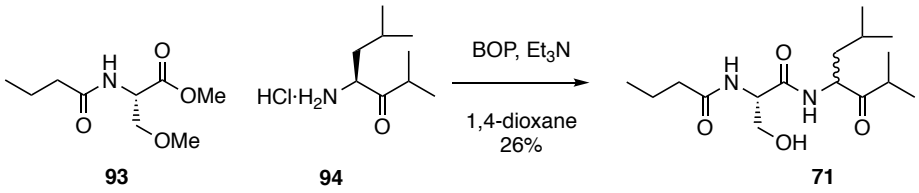
Figure 3.5: Mechanism of mono-addition of a Grignard reagent to a Weinreb amide via stable tetrahedral intermediate 119 that decomposes upon acidic work up.[74]

further improved by determining the concentration of the Grignard reagent via titration. However, attempts at optimisation were postponed until after it was determined whether **71** was a substrate for EpnF. For the final reaction to produce this

half of **71**, an acid-catalysed Boc deprotection was performed, producing **94** in a 55% yield (scheme 3.4).

3.1.2.4. Peptide coupling to produce isopropyl ketone substrate 71

The final step in the synthesis of **71** was the coupling of **94** and **93**. BOP was selected as the coupling reagent due to the high yield of 77% achieved previously for **96**. Although monitoring of the reaction mixture was attempted using thin layer chromatography (TLC) and MS, the polar nature of the product and reagents as well as the presence of readily ionisable species such as triethylamine rendered this challenging. In the absence of effective monitoring, the reaction was left for four hours, consistent with the procedure for **96**. However, this produced a much lower yield of 26% for **71**.



Scheme 3.5: Synthesis of isopropyl ketone 71 via BOP-mediated coupling.

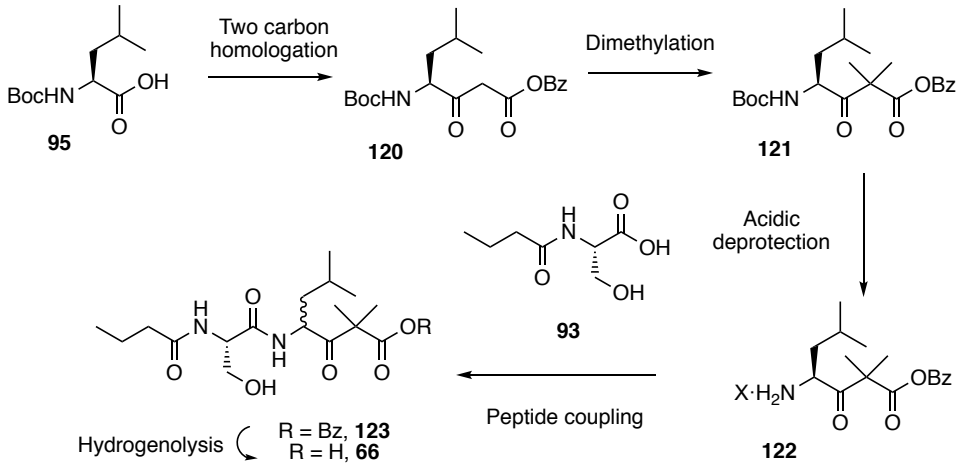
The yield of **71** could likely be improved by increasing the duration of the reaction, or by increasing the amount of base used, as reported for the final peptide coupling in the total synthesis of carmphycin A **15**.[22]

3.1.3 Synthesis of potential α-dimethyl-β-keto carboxylic acid substrate 66

3.1.3.1. Proposed synthetic route

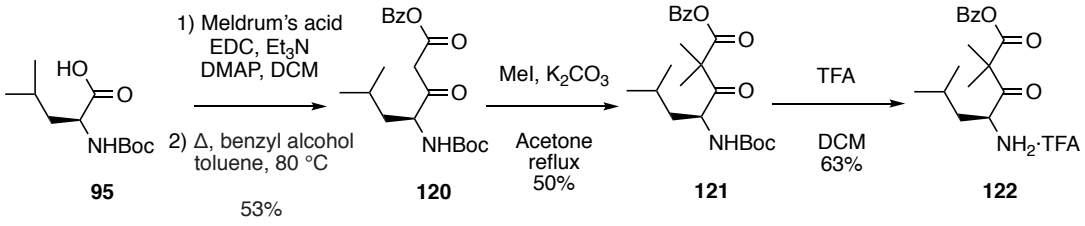
With peptide **93** available from the synthesis of **71**, a convergent synthesis was proposed for the production of α-dimethyl-β-keto carboxylic acid **66** (scheme 3.6). To produce the leucine derived α-dimethyl-β-keto benzyl ester **122**, coupling to

Meldrum's acid was proposed followed by thermolysis in the presence of benzyl alcohol to give **120**. Subsequent α-dimethylation using potassium carbonate and iodomethane in an S_N2 reaction would produce **121**.[77] As in the synthesis of **71**, Boc deprotection to give **122** and coupling would afford the α-dimethyl-β-keto benzyl ester **123**, which could be converted readily to the carboxylic acid **66** by removing the benzyl ester protecting group via hydrogenolysis over palladium on activated carbon.



Scheme 3.6: Proposed synthesis of putative substrate 66.

3.1.3.2. Synthesis of 122



Scheme 3.7: Synthesis of 122 by Meldrum's acid adduct thermolysis in the presence of benzyl alcohol, dimethylation with iodomethane in the presence of potassium carbonate, and Boc deprotection.[42]

The synthesis of benzyl ester **122** was completed by Dr Douglas Roberts, a postdoctoral research fellow in the Challis group with extensive experience of similar procedures (scheme 3.7), in order to complete the research more rapidly for publication.

The first steps in the synthesis of the α-dimethyl-β-keto benzyl ester **122** were to couple Meldrum's acid to *N*-BOC-L-leucine **95** using EDC **110** and DMAP **116** (scheme 3.7), and then heat the resulting adduct in the presence of benzyl alcohol to furnish the β-keto benzyl ester **120**. A commonly reported mechanism of thermolysis, after addition of alcohol to one of the carbonyl functionalities to give **126**, is concerted degradation driven by the evolution of carbon dioxide to provide a β-keto ester (figure 3.6).[77]

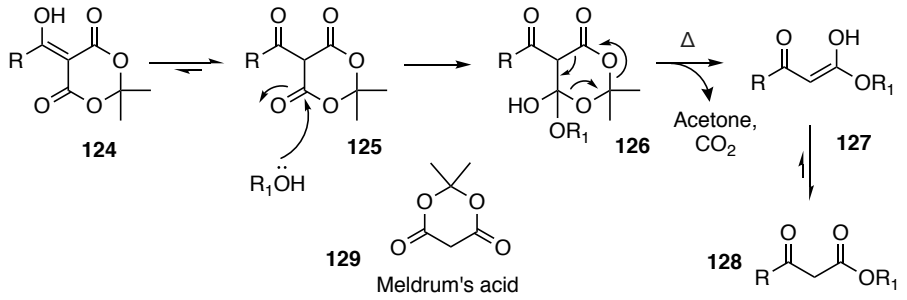


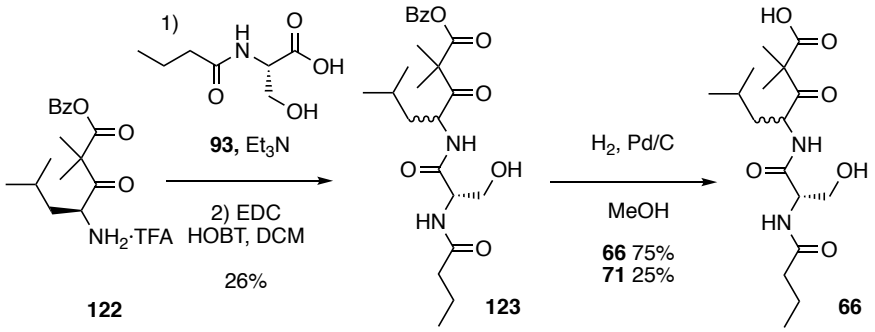
Figure 3.6: Proposed mechanism of thermolysis of a Meldrum's acid adduct in the presence of an alcohol to form β-keto esters 128.[77]

α-Dimethylation of **120** was accomplished by reaction with potassium carbonate and iodomethane, with the acidic α-position between the carbonyl functionalities readily deprotonated due to stabilization of the resultant anion by delocalization, yielding **121** via two sequential methylation reactions. Lastly, BOC-deprotection was achieved using trifluoroacetic acid to produce **122** (scheme 3.7).

3.1.3.3. Peptide coupling and hydrogenation to produce 66

The synthesis of **66** was completed via peptide coupling of **93** (section 3.1.2.2) to **122**, using EDC **110** with HOBT **115** as an additive. While the yield for this reaction was

low (26%) and resulted in a mix of diastereomers, additional optimization was not attempted due to the instability of **66**, which was formed via hydrogenolysis over palladium on activated carbon. Approximately 25% of **66** underwent decarboxylation during the course of this reaction to form ketone **71** (scheme 3.8).



Scheme 3.8: Synthesis of α-dimethyl-β-keto carboxylic acid 66 via EDC coupling and hydrogenolysis of the benzyl ester protecting group.

As previously discussed (section 1.3.2.1), α-dimethyl-β-keto acids are prone to spontaneous decarboxylation to form an enol, which readily tautomerises to the corresponding isopropyl ketone (figure 3.7). This was observed on purification of **66**, with isopropyl ketone **71** visible by MS and ^{1}H NMR analysis and easily identifiable due to being previously synthesized (section 3.1.2.4). The amount of **71** increased over

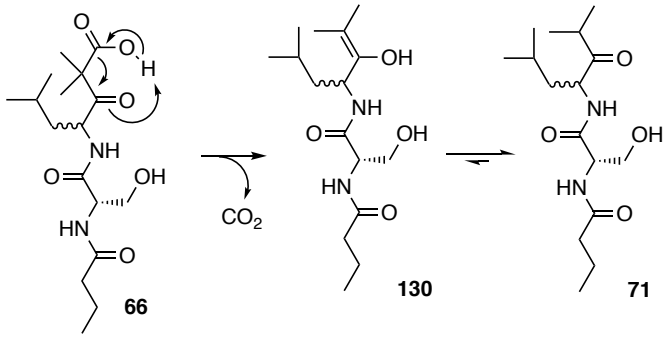


Figure 3.7: Proposed mechanism of decarboxylative degradation to convert the α-dimethyl-β-keto acid 66 to the corresponding isopropyl ketone 71.

time, indicative of spontaneous degradation rather than a process that is promoted by the conditions of the hydrogenolysis reaction.

To reduce the level of degradation observed in the hydrogenolysis reaction, an excess of Pd/C was used, facilitating rapid benzyl deprotection and minimizing the amount of time **66** spent in solution. This allowed **66** to be used for *in vitro* assays with EpnF.

3.2 Activity assays with purified recombinant EpnF and TmcI

3.2.1. EpnF assay with α-dimethyl-β-keto acid 66

The ability of EpnF to catalyse epoxidation of either **71** or **66** was examined by incubating each with purified recombinant His_6-EpnF for three hours. The enzyme was then precipitated with acetonitrile, and the supernatant was analysed by UHPLC-ESI-Q-TOF MS.

Despite screening a range of conditions, no formation of epoxyketone **78** was observed in assays containing **71**. Similarly, no activity was observed when this substrate was incubated with TmcI, NADPH, spinach ferredoxin and ferredoxin reductase, consistent with the fact that the macyranone A **20** biosynthetic gene cluster contains an *epnF* homologue, but no *tmcI* homologue.[25] This suggested that EpnF does not function as a conventional ACAD (figure. 1.14A), and may instead catalyse successive decarboxylation, desaturation and epoxidation of an *α*-dimethyl-*β*-keto acid (figures 3.8, 5.4).[56]

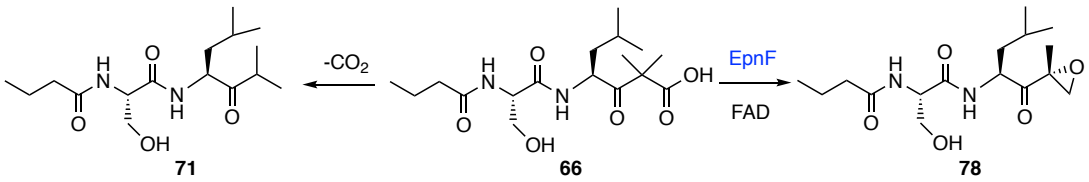


Figure 3.8: Products observed by UHPLC-ESI-Q-TOF MS when 66 was incubated with His_6-EpnF.

Consistent with this hypothesis, a compound with a molecular formula corresponding to epoxyketone **78** was observed when **66** was incubated with His_6-EpnF (figures 3.8 and 3.9). This compound was absent in a negative control containing heat denatured His_6-EpnF. Interestingly, no cofactors were needed apart from the enzyme-bound flavin (section 2.2.3.2). The production of **78** was accompanied by consumption of **66** and decarboxylation of **66** to form **71** was observed in both the enzymatic reaction and the negative control (figures 3.8 and 3.9). This strongly indicated that α-dimethyl-β-keto acids are the substrate for this novel class of flavin-dependent enzymes. To the best of the author's knowledge, the reaction proceeds *via* a previously uncharacterized decarboxylation, desaturation and epoxidation catalytic mechanism, which is discussed in detail in chapter 5.

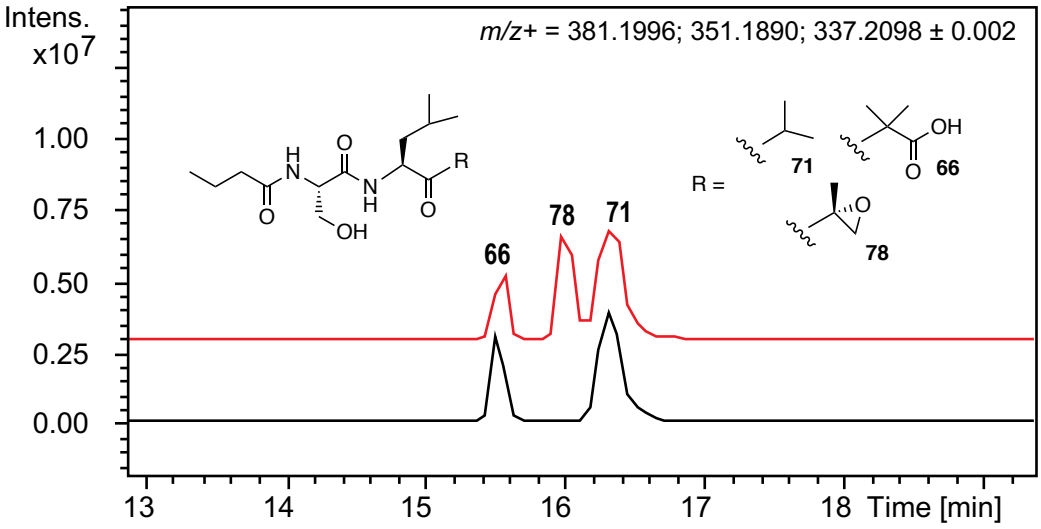


Figure 3.9. EICs at *m/z* = 381.20, 351.19, 337.21 corresponding to $[M + Na]^+$ for 66, 78 and 71 respectively from UHPLC-ESI-Q-TOF MS analyses of assays containing 66 and native (top trace) or heat-denatured (bottom trace) His_6-EpnF.

3.2.2 Coupled assay of EpnF and TmcI activity

A remaining undetermined aspect of both TMC-86A and eponemycin biosynthesis was the role of the homologous P450 enzymes TmcI and EpnI. Both natural products have a α-hydroxymethyl group attached to one of the carbon atoms of the epoxide. It

was thus hypothesized that TmcI/EpnI could be responsible for forming this group via a hydroxylation reaction.

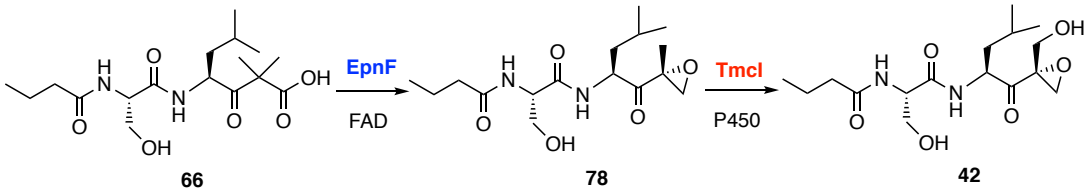


Figure 3.10: Reactions catalysed by EpnF and TmcI with 66.

To elucidate the role of TmcI, an *in vitro* assay containing recombinant His_6-EpnF and His_6-TmcI was conducted. In order for cytochromes P450 to catalyse monooxgenation reactions, they undergo reduction through an electron transport chain as previously discussed (section 1.3.2.3) In *Streptomyces* species, the electrons are usually supplied by NAD(P)H via the redox partners ferredoxin, ferredoxin reductase, but commercially available ferredoxin and ferredoxin reductase from *Spinacia oleracea* can be utilized as surrogates.[61]

66 was therefore incubated with His_6-EpnF, His_6-TmcI, ferredoxin and ferredoxin reductase from spinach, and NADPH. Following acetonitrile precipitation of the proteins, UHPLC-ESI-Q-TOF MS analysis of the supernatant revealed the production of hydroxylated epoxyketone **42** (figure 3.10). No production of this compound was observed in negative control reactions containg boiled His_6-EpnF or His_6-TmcI (figure 3.11). These data show that TmcI is able to hydroxylate **78**, but not **66**.

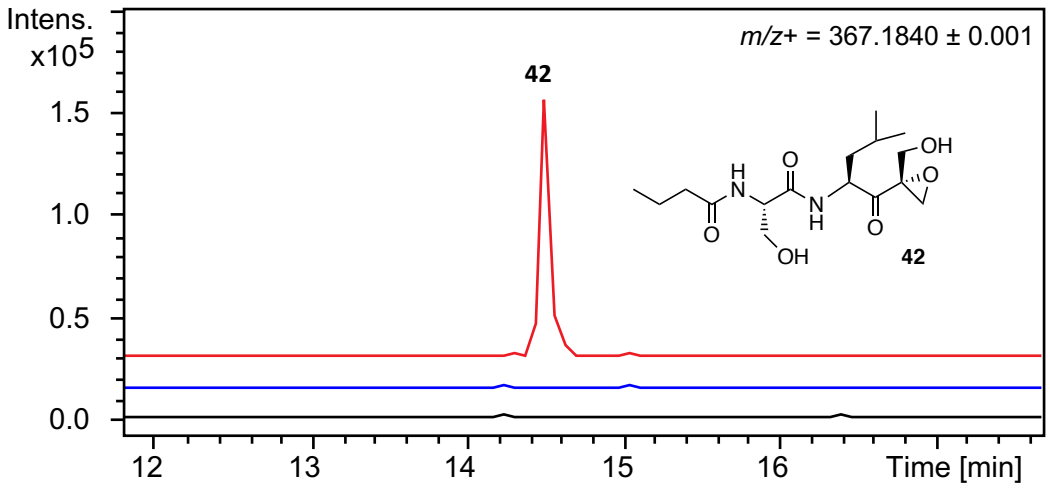


Figure 3.11: EICs at *m/z* = 367.18, corresponding to [M + Na]⁺ for 42, from UHPLC-ESI-Q-TOF MS analyses of assays containing 66, NADPH, ferredoxin and ferredoxin reductase, and either native His_6-EpnF and His_6-TmcI (top trace), heat denatured His_6-EpnF and native His_6-TmcI (middle trace), or native His_6-EpnF and heat denatured His_6-TmcI (bottom trace).

3.2.3 MS/MS analysis of products

Due to the inherent instability of the substrate **66**, a larger scale assay to isolate epoxyketone **78** in sufficient quantity for NMR spectroscopic analysis proved impossible. To further characterise the products of the EpnF and TmcI catalysed reactions they were subjected to MS/MS analyses by Dr. D. Zabala and the fragment ions produced were compared to those observed for TMC-86A. Signature sets of fragment ions were observed across the three species (figure 3.12), consistent with the proposed structures of **78** and **42**.[42]

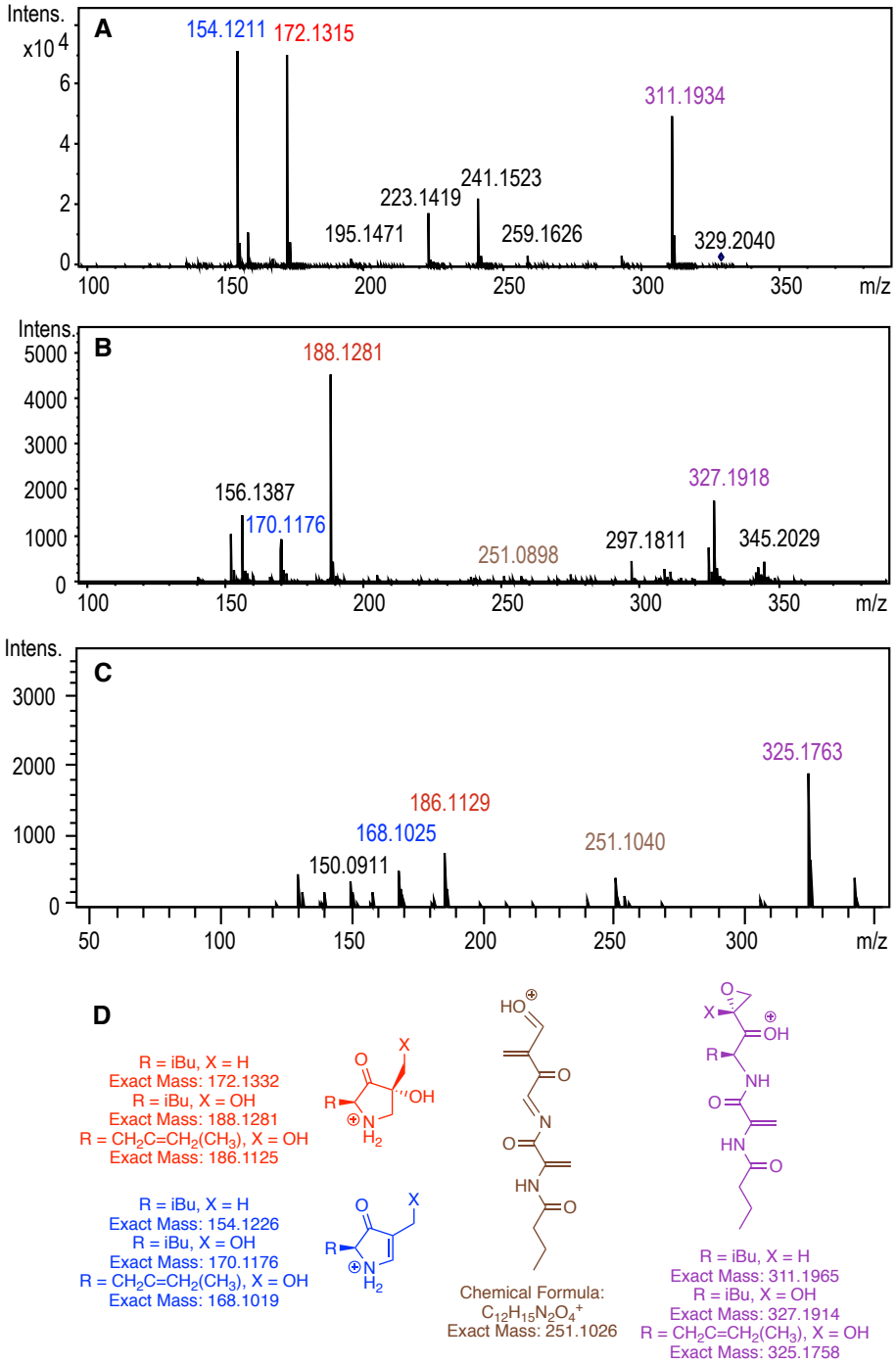


Figure 3.12 A: MS/MS spectrum for the *m/z* = 329.20 ion, corresponding to $[M + H]^+$ for 78. B: MS/MS spectrum for the *m/z* = 345.20 ion, corresponding to $[M + H]^+$ for 42. C: MS/MS spectrum for the *m/z* = 343.19 ion, corresponding to $[M + H]^+$ for TMC-86A. D: Proposed structures of observed fragment ions from MS/MS fragmentation, from dehydration and cleavage of the amide bond followed by intramolecular cyclisation.

3.2.4 Reaction of EpnF assay product with *N*-acetylcysteamine

Derivatization of the epoxyketone assay product **78** by nucleophilic ring-opening of the epoxide was investigated, to provide additional evidence for the presence of the pharmacophore. Thiols are slow to react with epoxide functionalities, with literature citing the necessity of 37 °C, a five-fold excess of *N*-acetylcysteamine and a time scale of over 10 hours to drive an aqueous reaction to completion.[78] In the instance of **78**, the presence of the electron withdrawing α-keto group was proposed to activate the epoxide towards nucleophilic attack (figure 3.13), as observed for the reaction of epoxyketones with the 20S proteasome (section 1.2.1).

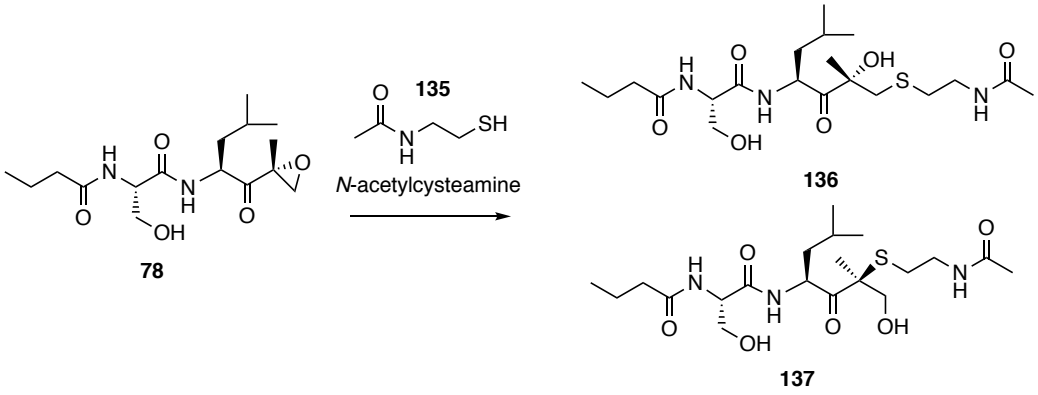


Figure 3.13: Adducts 136 and 137 proposed to result from opening of the epoxide in the product of the EpnF-catalysed oxidation of 66 with *N*-acetylcysteamine.

The substrate **66** was incubated with His_6-EpnF in the presence of an excess of *N*-acetylcysteamine for three hours at room temperature. Upon acetonitrile precipitation of the protein, UHPLC-ESI-Q-TOF MS analysis of the supernatant revealed ion species with an *m/z* corresponding to the *N*-acetylcysteamine adduct of epoxyketone **78** (figure 3.14). As there was a shoulder peak observed in the chromatogram trace, it is possible that both **136** and **137** were formed from reaction with the thiol. Neither of these species were present in the negative control using heat denatured EpnF. These data are consistent with the proposal that **78** contains an epoxide.

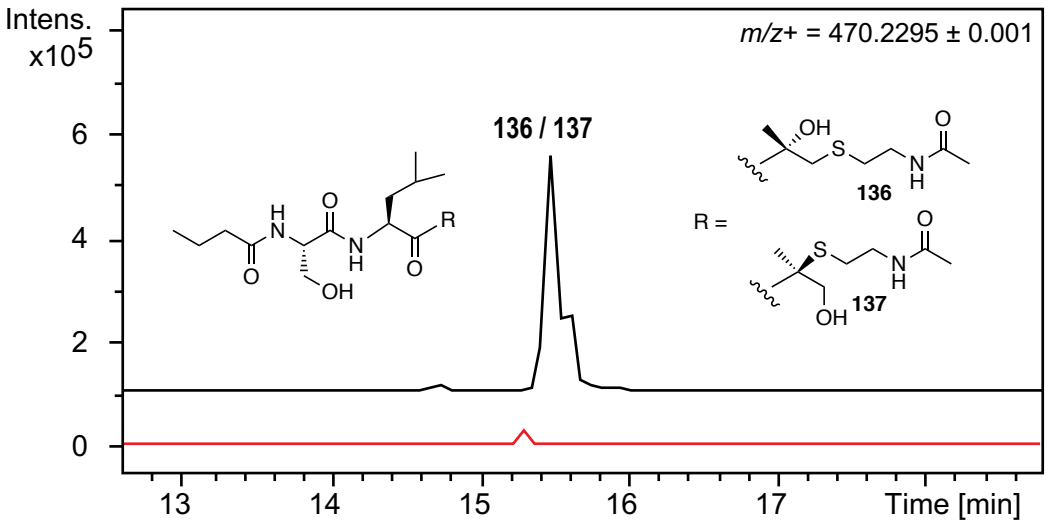


Figure 3.14: EICs at *m/z* = 470.23, corresponding to [M + Na]⁺ for 136/137, from UHPLC-ESI-Q-TOF MS analysis of reaction mixture containing 66, an excess of *N*-acetylcysteamine and either native (top trace), or heat denatured (bottom trace) His_6-EpnF.

3.3 Conclusions

The synthesis of both the isopropyl ketone **71** and α-dimethyl-β-keto carboxylic acid **66** was accomplished, and each was incubated with EpnF. The epoxyketone **78** was only produced in assays containing **66** and not **71**, supporting the hypothesis that α-dimethyl-β-keto carboxylic acids are the true substrates of this unusual class of flavin-dependent decarboxylase-dehydrogenase-monooxygenases.[42] Our results were independently corroborated by Zhang and coworkers, who showed that co-expression of EpnG, EpnH and EpnF in *E. coli* results in the production of epoxyketones upon feeding of appropriate *N*-acetylcysteamine thioesters.[79]

TmcI was shown to function as a hydroxylase by incubating **66** with both EpnF and TmcI, as well as ferredoxin and ferredoxin reductase from spinach, and NADPH. This produced a compound with a molecular formula corresponding to the hydroxylated epoxyketone **42**. Although Zhang and coworkers co-expressed the homologous enzyme EpnI from the eponemycin biosynthetic gene cluster with EpnG, EpnH and EpnF, they did not observe formation of an analogous hydroxylated product.

Presumably, this is due to poor compatibility of *E. coli* redox proteins with cytochrome P450 enzymes from *Streptomyces* species.[61]

The pattern of fragment ions observed in MS/MS analyses of **78** and **42** were consistent with those observed for TMC-86A, providing support for the proposed structures of the products of the EpnF and TmcI-catalysed reactions. The addition of *N*-acetylcysteamine to **78** provided additional evidence that it contains an epoxyketone moiety.

Decarboxylation of **66** to form the isopropyl ketone **71** presented a significant obstacle to further characterisation of the EpnF-catalysed reaction. In an attempt to circumvent this problem, methyl ester analogues of **66** were synthesised and various methods for their hydrolysis to the corresponding β-keto acids were investigated, as detailed in Chapter 4.

Chapter 4: β-Keto methyl esters as stable precursors to EpnF substrates

4.1 Synthesis of leucine derived α-dimethyl-β-keto methyl ester

4.1.1 Rationale

As previously discussed (sections 1.3.2.1 and 3.1.2.3), many α-dimethyl-β-keto carboxylic acids are metastable, and are prone to spontaneous decarboxylation and subsequent tautomerization to form ketones. To make the EpnF-catalysed reaction more efficient and to provide scalable improved methods for accessing its α-dimethyl-β-keto acid substrate, an alternate substrate precursor was needed. Unlike α-dimethyl-β-keto carboxylic acids, α-dimethyl-β-keto methyl esters **138** are not prone to spontaneous degradation to ketones under ambient conditions. Instead Krapcho decarboxylation (≥150 °C, LiCl) is required for the analogous reaction to occur (figure 4.1A).[80] Therefore the synthesis of **139**, the methyl ester of **66**, was targeted with the intention of investigating mild methods for converting it to the corresponding carboxylic acid (figure 4.1B).

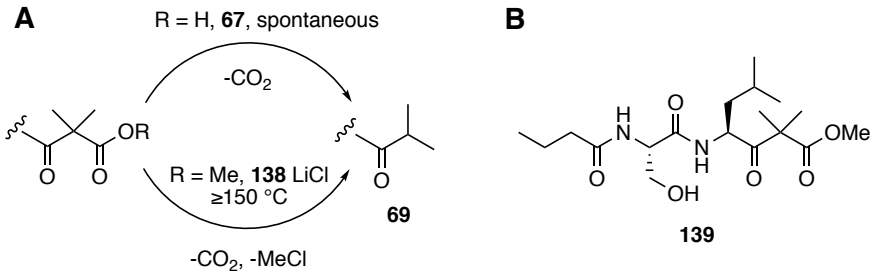


Figure 4.1. A: Conditions required to transform an α-dimethyl-β-keto carboxylic acid or its methyl ester into an isopropyl ketone 69. B: Structure of methyl ester precursor 139 of the EpnF substrate 66 targeted for synthesis.

As the carboxylic acid **66** had been demonstrated to be a substrate for EpnF, the methyl ester **139** would be an appropriate precursor, provided a mild method for hydrolysing it could be identified. One possibility was to use a promiscuous hydrolase such as commercially available pig liver esterase (PLE) to generate the acid *in situ*.

PLE is a serine hydrolase which uses a catalytic triad of serine, histidine and aspartate residues. This triad enables nucleophilic attack of an ester substrate by the serine hydroxyl group, forming an acyl-ester intermediate, which undergoes hydrolytic cleavage.[81] Preparations of PLE are mixtures of isoenzymes with different substrate specificities, with each isoenzyme composed of a trimer of different enzyme subunits. While this heterogeneity can be problematic when using PLE for enantioselective catalysis in organic synthesis, it helps to ensure substrate promiscuity in other applications.[82] In lieu of an esterase with high specificity for **139**, PLE seemed an appropriate enzyme with which to begin investigation into *in situ* substrate precursor hydrolysis (figure 4.2).

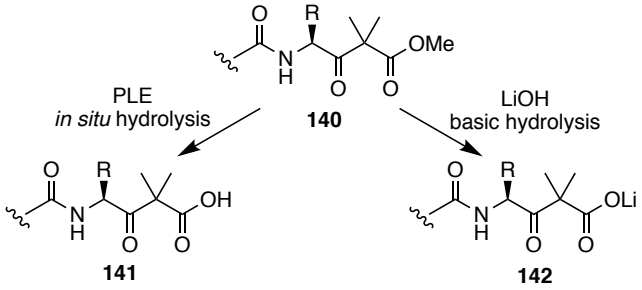


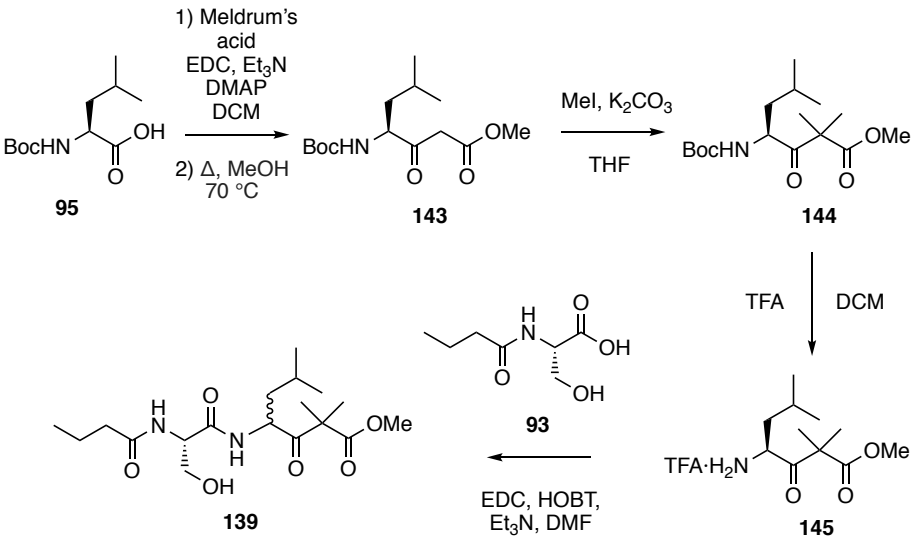
Figure 4.2: The chemical and biological strategies for preparation of EpnF substrates.

As discussed previously, α-dimethyl-β-keto carboxylic acids **141** lacking substituents that stabilize the enol tautomer are prone to decarboxylate through a cyclic transition state due to intramolecular hydrogen bonding (section 1.3.2.1). The lithium salt **142** of such carboxylic acids lack the ability to form this intramolecular hydrogen bond, and as such may prove to be more stable (figure 4.2). Indeed, stable lithium *α*-alkyl-*β*-keto carboxylates have been reported in the literature.[83]

4.1.2 Synthetic plan

The previous synthesis of **66** involved the use of a benzyl ester protecting group on the carboxylic acid, which was removed by hydrogenolysis in the final step.[42] It was therefore reasoned that the existing synthesis could be adapted by thermolysis of a

leucine-derived Meldrum's acid adduct in the presence of methanol instead of benzyl alcohol to produce methyl ester **143**. Following this, completion of the same series of steps as used previously would provide the desired α-dimethyl-β-keto methyl ester **139** (scheme 4.1).



Scheme 4.1: Proposed synthesis of methyl ester substrate precursor 139.

4.1.3 Synthesis of 144

N-BOC-L-Leucine **95** was first coupled to Meldrum's acid **129** using EDC **110** and DMAP **116**, followed by thermolysis in the presence of methanol to afford **143**.[42] In contrast to the previously employed procedure to produce **120** (section 3.1.3.2), monitoring by TLC revealed BOC-L-leucine **95** could still be observed after 27 hours. Therefore additional equivalents of DMAP **116** and EDC **110** (0.4 equivalents, increasing the total of each from 1.1 to 1.5) were added to drive the reaction to completion. This afforded a crude yield of 79% for the Meldrum's acid adduct.

In the second step of the reaction methanol was used as both solvent and reactant under refluxing conditions. Methanol reacted with the Meldrum's acid adduct as previously described for **120** (figure 3.6). Following purification **143** was obtained in 31% yield.

Since this was lower than the yield of the benzyl ester **120** (51%), and this reaction was the first step of the synthesis, optimization was desirable.[42] Thermolysis of Meldrum's adducts of Boc-protected amino acids are reported to form lactam byproducts (putatively resulting from nucleophilic attack and cyclisation by the amide nitrogen) under refluxing conditions.[84] While only **143** was retained and characterized from silica gel chromatography, ions corresponding to a lactam byproduct were observed by MS analysis of the reaction mixture. The low yield in conjunction with this possibility of byproduct formation prompted a change of strategy.

An alternative coupling partner, methyl potassium malonate, was investigated. This forms β-keto esters with carbonyl diimidazole (CDI) **146** activated carboxylic acids.[85,86] Similarly to the BOP, DCC and EDC reagents mentioned previously (sections 3.1.2.2 and 3.1.2.3), CDI activates carboxylic acids towards nucleophilic addition-elimination by formation of adduct **150** with a good leaving group (figure 4.3), and drives the reaction through evolution of carbon dioxide.[76]

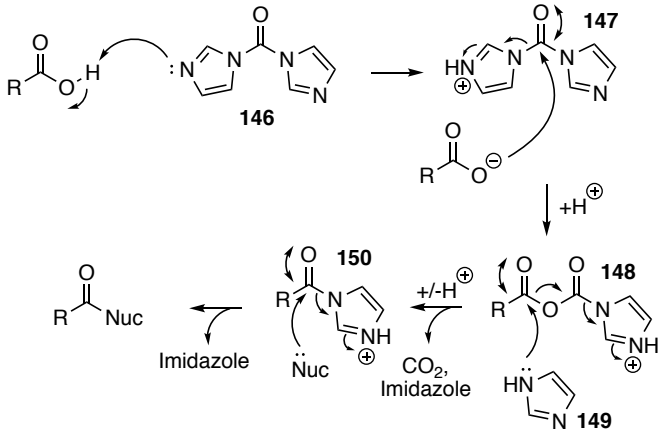
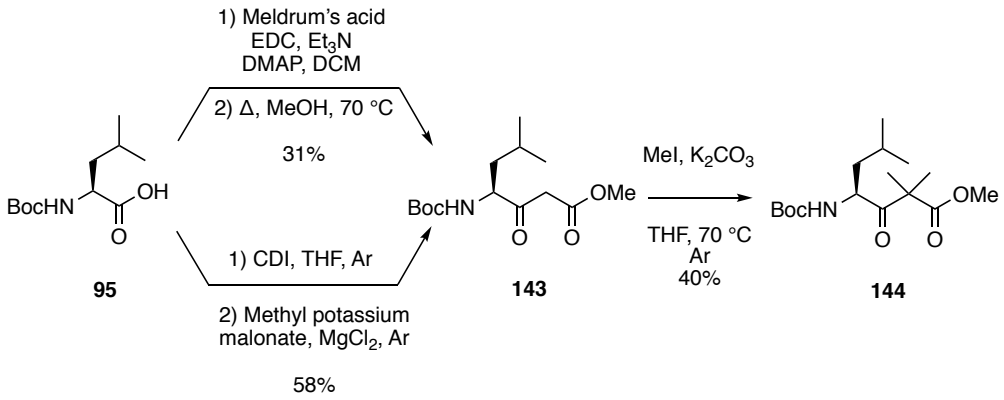


Figure 4.3: Mechanism of CDI-mediated coupling of a carboxylic acid with a generic nucleophile.[76]

This new approach offered several practical advantages over Meldrum's acid. Although two steps are still required, the first step to form the reactive imidazole adduct takes two hours, rather than the 24 hours required for Meldrum's acid adduct formation. Additionally, **143** could be isolated from the reaction using mild acidic and

basic washes, to remove imidazole and unreacted magnesium-malonate half ester salt, without the need for silica gel chromatography. This lead to an increase in yield from 31% to 58% for **143** (scheme 4.2).

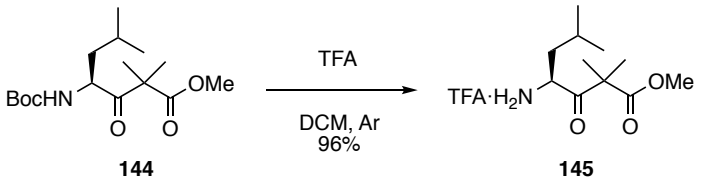


Scheme 4.2: Alternate synthetic routes to β-keto methyl ester 143, and dimethylation to produce α-dimethyl-β-keto methyl ester 144.

To form the α-dimethyl-β-keto methyl ester **144**, the carbon between the two carbonyl groups was methylated using iodomethane and potassium carbonate, as described previously (section 3.1.3.2), in a 40% yield (scheme 4.2).

4.1.4 Synthesis of 145

Initially the same conditions used for BOC removal of **121** (stirring in 10% TFA in DCM over 17 hours) were applied to remove the BOC group from **144**.[42] However these conditions did not yield the desired product, presumably because the methyl ester of **144** was more prone to acid hydrolysis than the benzyl ester of **121**. By changing to completely anhydrous conditions and closely monitoring the reaction by TLC, the deprotection was found to go to completion in one hour, providing **145** in 96% yield (scheme 4.3), an improvement on the 63% yield obtained for the benzyl ester **122**.

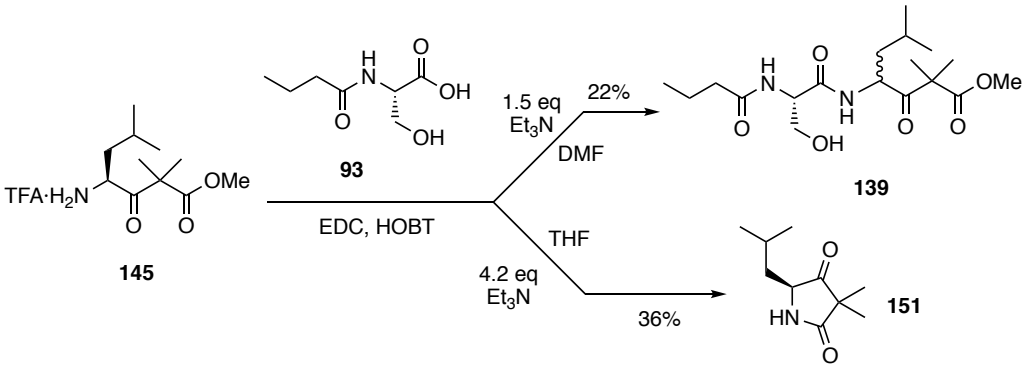


Scheme 4.3: Deprotection of α-dimethyl-β-keto methyl ester 144 using trifluoroacetic acid to give amine salt 145.

4.1.5 Coupling of 93 and 145

The final stage of the proposed synthesis was the coupling of carboxylic acid **93** to TFA salt **145**, using HOBT **115**, EDC **110** and triethylamine, as previously described (section 3.1.2.3). Using 1.5 equivalents of Et_3N, a yield of 22% was obtained (scheme 4.4), giving an overall yield for the convergent synthesis of 4.9%.

Other amide bond coupling reactions had been shown to produce higher yields using 4.2 equivalents of triethylamine and THF instead of DMF as the solvent (section 4.2.2). Ordinarily excess base can assist such coupling reactions by ensuring that carboxylic acids and ammonium salts are deprotonated. However, in this instance the amino group of **145** underwent intramolecular addition to the carbonyl group of the methyl ester, resulting in formation of lactam **151** as the major product (scheme 4.4).



Scheme 4.4: Attempted coupling of 145 and 93, and the major products observed in each case.

This result is likely due to the Thorpe-Ingold (or gem-dimethyl) effect, whereby increasing the number of substituents on a tetrahedral carbon centre favours intramolecular reactions within that molecule due to a reduction in the bond angle between reacting groups.[87,88]

Briefly, an alternate linear synthesis of **139** was attempted, however the coupling of *N*-butyryl-L-seryl-L-leucine **152** with methyl potassium malonate failed to produce any of the desired β-keto methyl ester **153** (figure 4.4). As this procedure had already produced **143** with *N*-BOC-L-leucine **95** (scheme 4.2), it was reasoned that the hydroxyl group of serine interferes with the coupling reaction. Protection of this hydroxyl group was consequently deemed to be required for further synthetic efforts (section 4.2).

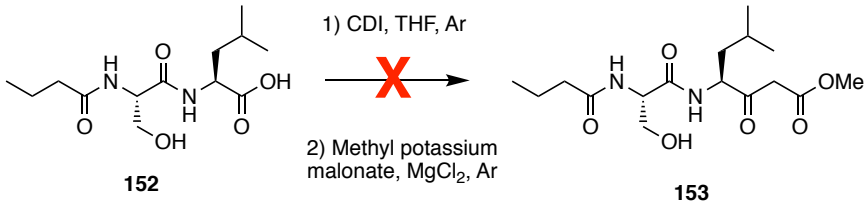


Figure 4.4: Unsuccessful CDI-mediated coupling of 152 and methyl potassium malonate.

4.1.6 Pig liver esterase and EpnF in vitro assays with 139

As only a modest amount of **139** was obtained via the synthetic route outlined above, the stability of the carboxylate salt resulting from base-promoted hydrolysis of the methyl ester was not investigated at this stage (section 4.2.6). However, the *in situ* hydrolysis of the ester using PLE was explored (figure 4.5).

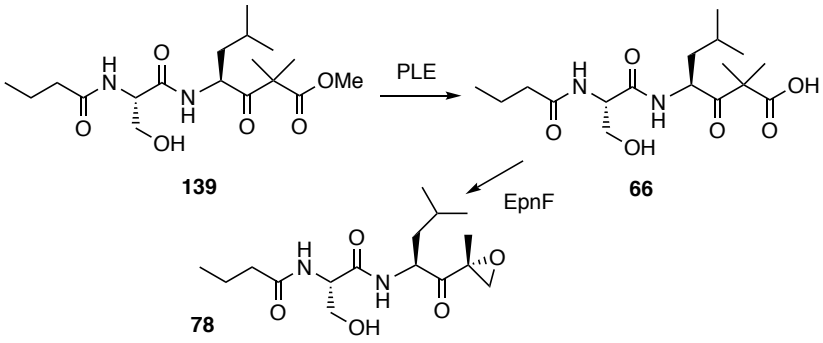


Figure 4.5: Proposed method for production and decarboxylation-dehydrogenation-epoxidation of 66 using PLE and EpnF in a single reaction vessel.

The ability of PLE to catalyse the hydrolysis of **139** was first investigated by incubating it with PLE at room temperature for three hours at pH 8 (section 7.4.4). After precipitation of the enzyme with acetonitrile, UHPLC-ESI-Q-TOF MS analysis of the supernatant confirmed production of **66**, which was absent with heat denatured PLE (figure 4.6). While a low intensity peak was observed in both reactions

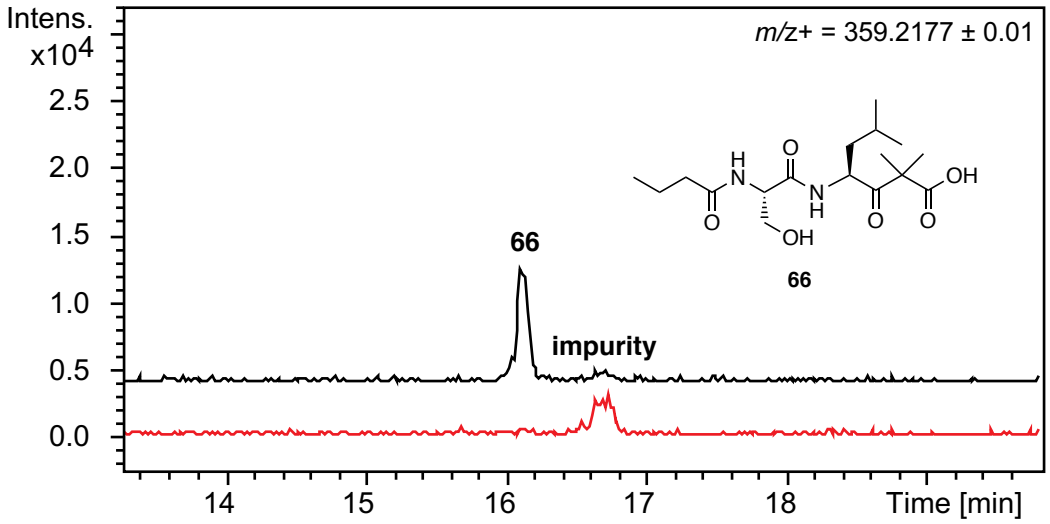


Figure 4.6. EICs at *m/z* = 359.22, corresponding to $[M + H]^+$ for 66, from UHPLC-ESI-Q-TOF MS analyses of assays containing 139 and native (top trace) or heat-denatured (bottom trace) PLE.

(approximately at 16.6 minutes), the corresponding ion did not have a molecular formula consistent with **66** by HR-MS analysis, and was therefore considered an unrelated impurity.

A corresponding decrease in the amount of **139** was also observed in the PLE-catalysed reaction relative to the negative control (figure 4.7).

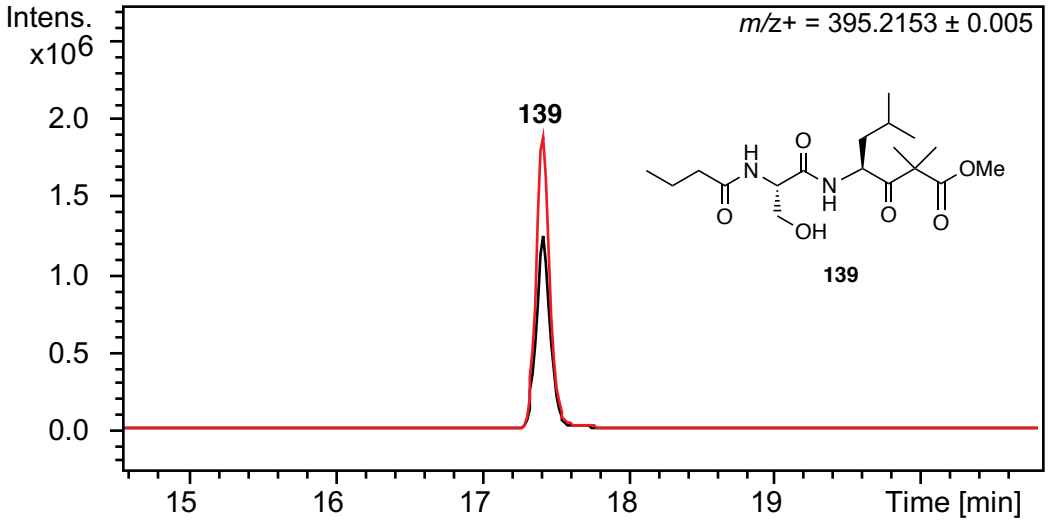


Figure 4.7: EICs at *m/z* = 395.22, corresponding to [M + Na]⁺ for 139, from UHPLC-ESI-Q-TOF MS analyses of assays containing 139 and native (black trace) or heat-denatured (red trace) PLE.

Having confirmed that PLE can hydrolyze **139**, the hydrolysis and subsequent decarboxylation-dehydorgenation-monooxygenation of **139** using PLE and purified recombinant His_6-EpnF in a single vessel was investigated next. The epoxyketone **78** was only produced in significant quantities when both PLE and EpnF were present (figure 4.7). A very small amount of **78** was also produced in the control reaction lacking PLE, presumbaly due to uncatalysed ester hydrolysis (figure 4.8). A negative control in which PLE and heat-denatured EpnF were present failed to produce any **78**.

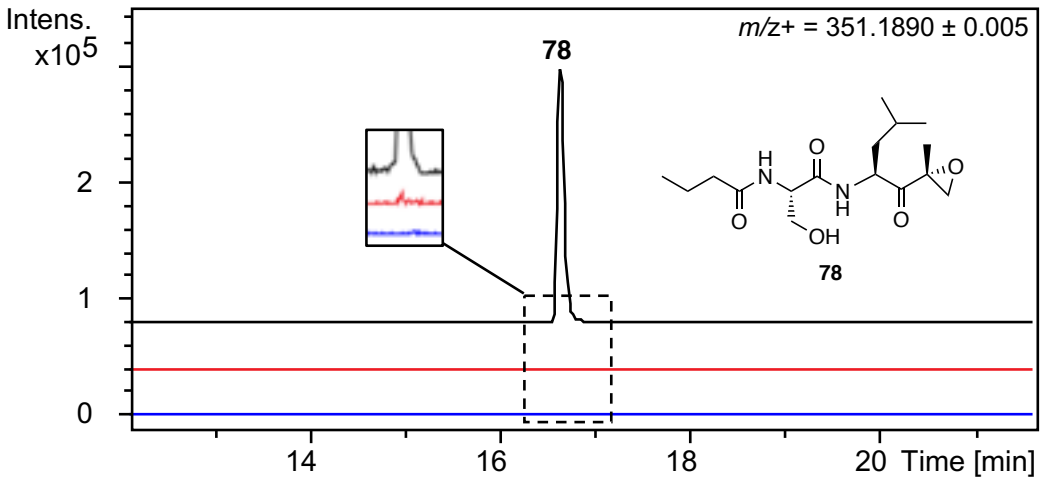


Figure 4.8: EICs at *m/z* = 351.19, corresponding to [M + Na]⁺ for 78, from UHPLC-ESI-Q-TOF MS analyses of assays containing 139 and either native His₆-EpnF and PLE (top trace), native His₆-EpnF (middle trace), or native PLE and heat-denatured His₆-EpnF (bottom trace).

This result provides an effective proof of concept for the use of α-dimethyl-β-keto methyl esters as stable precursors to the corresponding carboxylic acids and their one pot conversion to epoxyketones using EpnF in combination with an esterase.

4.2 Preliminary investigation into EpnF substrate tolerance

4.2.1 Synthesis of substrate analogue 154

As **66** had previously been shown to be a substrate for EpnF, and *in situ* hydrolysis of the corresponding methyl ester **139** using PLE proved to be an effective strategy for producing epoxyketone **78**, it was decided to investigate whether analogues of **139** with different amino acid side chains can be converted to epoxyketones using PLE and EpnF. While most naturally occurring epoxyketones are derived from either L-leucine or 4,5-dehydro-L-leucine, clarepoxcin A and tryptopeptin A incorporate L-alanine and L-tryptophan respectively (figure 4.9A).[23,24] An analogue of **139**, containing an alanine-derived β-keto ester, was thus targeted for synthesis (figure 4.9B).

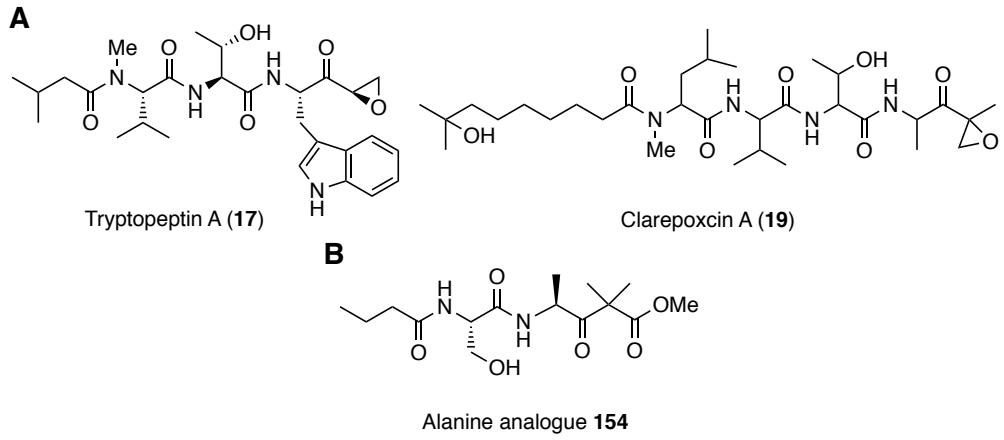
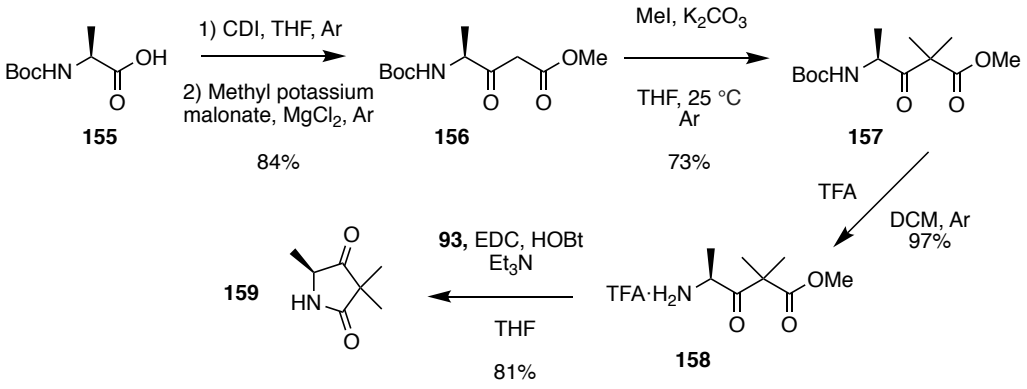


Figure 4.9. A: Structures of clarepoxcin A and tryptopeptin A, epoxyketones obtained from heterologous expression of a soil metagenomic library and *Streptomyces* sp. KUSC-G11, respectively.[23,24] B: Structure of alanine-derived analogue 154 of the substrate 139 used in the PLE/EpnF coupled assays.

4.2.1.1 Attempted adaptation of convergent synthetic route

Concurrently with synthesis of **139** (sections 4.1.3-4.1.5), the same procedures were applied to *N*-BOC-L-alanine in an attempt to produce **154** (scheme 4.5). Yields for these procedures increased significantly compared to previously obtained yields for *N*-BOC-L-leucine. Meldrum's acid coupling to methyl potassium malonate yielded 84% of **156** compared to 58% for **143**, and also an increase from 69% recorded in the literature for **156**.[86] Under milder reaction conditions, **157** was produced in a yield of 58% using iodomethane and potassium carbonate, compared to the 40% yield of **144**. Deprotection of the Boc group using TFA produced **158** in a yield of 97%, consistent with the 96% yield of **145**.

In the final coupling step to **93**, the same issue of lactamization occurred, producing the byproduct **159** in a yield of 81% (scheme 4.5), with none of the desired methyl ester **154**. This provided additional confirmation of the inherent problems of convergent synthesis for this application.

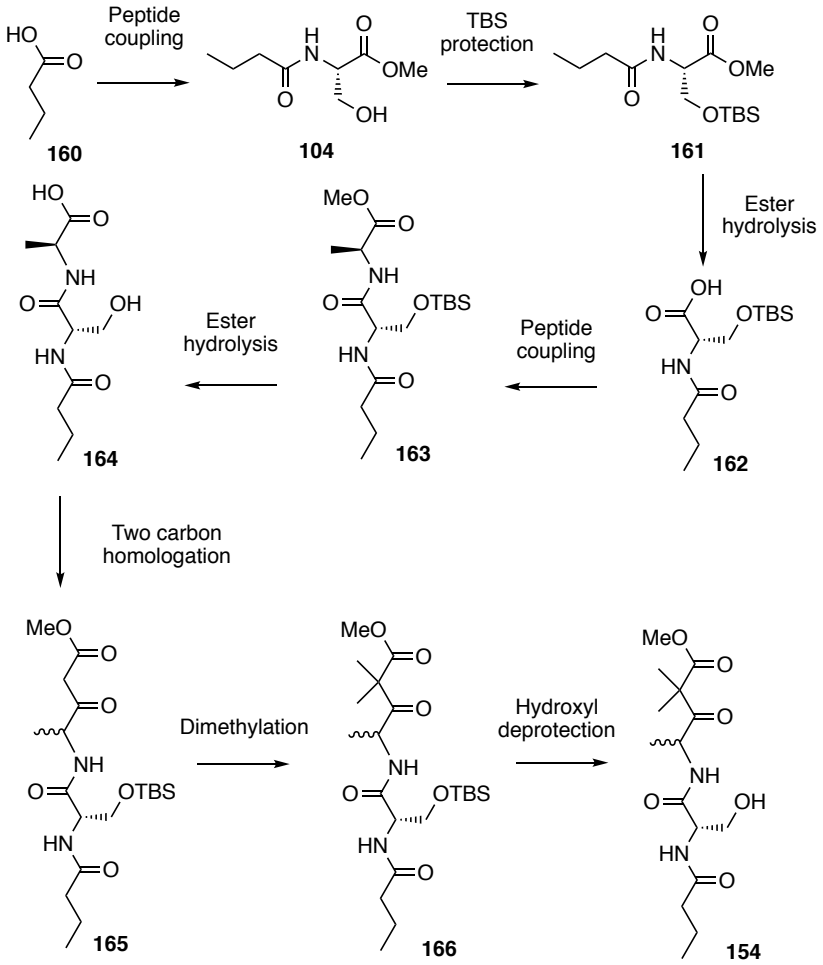


Scheme 4.5: Attempted convergent synthesis of 154.

4.2.2 Proposed linear synthesis

Convergent synthesis of α-dimethyl-β-keto methyl ester compounds had so far proved to be problematic, yielding both **151** and **159** as lactam byproducts (sections 4.1.5 and 4.2.1.1). Additionally, a brief attempt at a linear synthesis of **139** had failed when trying to couple *N*-butyryl-L-seryl-L-leucine **152** to methyl potassium malonate using CDI (figure 4.4). As malonate coupling had been successful for Boc-protected leucine, it was reasoned that the hydroxyl side chain of serine must be interfering with the reaction.

Therefore, a new linear synthesis of **154** was devised. Protection of the **104** serine hydroxyl group with *tert*-butyldimethylsilyl (TBS) chloride would give **161**, which was anticipated to tolerate the basic conditions used in the peptide coupling and ester hydrolysis reactions used to produce **164** (scheme 4.6). Two carbon homologation of **164** using CDI and methyl potassium malonate (sections 4.1.3 and 4.2.1.1), followed by dimethylation with iodomethane and potassium carbonate, and removal of the TBS protecting group, would give the desired methyl ester **154**.

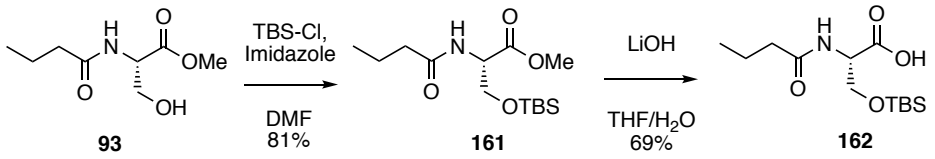


Scheme 4.6: Linear route devised for the synthesis of alanine substrate precursor analogue 154.

While *N* to *C*-terminal peptide synthesis suffers from the problem of oxazolone formation, as previously discussed (figure 3.3), the observed lactamization from intramolecular addition of the amino to α-dimethyl-β-keto methyl ester in attempted peptide coupling makes the reverse strategy impractical.[76] In addition, two carbon homologation of dipeptides using CDI **146** and methyl potassium malonate has been reported to result in lower levels of α-carbon epimerization compared to other strategies, such as coupling with Meldrum's acid **129** and subsequent thermolysis in the presence of methanol.[89]

4.2.1 Synthesis of *O*-TBS-*N*-butanoyl-L-serine 162

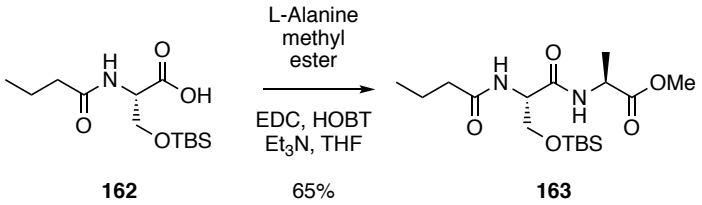
Protection of the side chain of previously synthesised **104** (section 3.1.2.2) was accomplished by nucleophilic substitution, using imidazole to activate the hydroxyl side chain and TBS-chloride. The method used was adapted from a literature procedure, in which all the reactants were added at the start of the reaction, giving **161** in a 53% yield.[90] By changing to dropwise addition of TBS-chloride over an hour, this yield was increased to 81%. Ester hydrolysis was completed using lithium hydroxide to produce **162** (scheme 4.7).



Scheme 4.7: Synthesis of TBS protected 161, with yield optimisation through dropwise addition, and subsequent methyl ester hydrolysis to provide 162.

4.2.2 Coupling of *O*-TBS-*N*-butanoyl-L-serine 162 and L-alanine methyl ester

Coupling of **162** and L-alanine methyl ester was attempted using HOBt **115**, EDC **110** and triethylamine using equivalents that had previously given high yields (1.0 equivalents of **162** and L-alanine methyl ester, 1.2 equivalents of HOBt **115**, 2.0 equivalents of EDC **110** and 4.2 equivalents of triethylamine). A workup for the reaction was avoided, as it had been observed that similarly hydrophilic compounds produced in the attempted L-leucine linear synthesis (figure 4.4) were extracted in the aqueous layer. Instead methyl ester **163** was purified by silica gel chromatography, providing a yield of 56%. In order to optimize the reaction further, the amount of HOBt **115** was increased to 2.4 equivalents, increasing the yield to 65% (scheme 4.8).



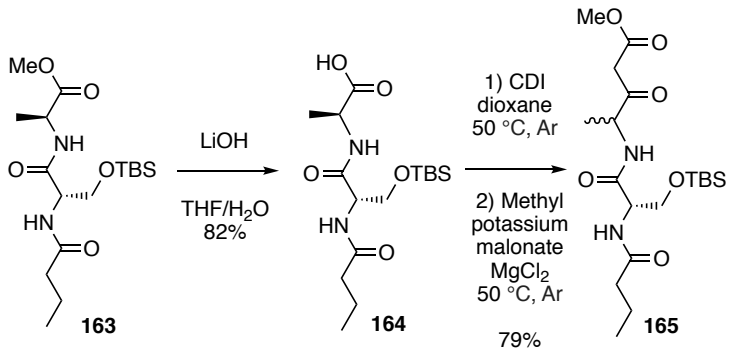
Scheme 4.8: Synthesis of methyl ester 163.

4.2.3 Hydrolysis of methyl *O*-TBS-*N*-butyryl-L-seryl-L-alaninate 163 and subsequent malonate coupling

To hydrolyze methyl ester **163**, lithium hydroxide was again employed. The reaction time was reduced, and strong acid was avoided during the workup to avoid TBS deprotection. Carboxylic acid **164** was produced in 80% yield.

The β-keto methyl ester **165** was prepared via the CDI **146**, methyl potassium malonate and magnesium chloride procedure previously established (sections 4.1.3 and 4.2.1.1). However, in this instance initial yields were low (~7%). **164** was recovered from the reactions, and resubmitted to the coupling procedure. This resulted in improved yields, ranging from 24% to 40% of **165**.

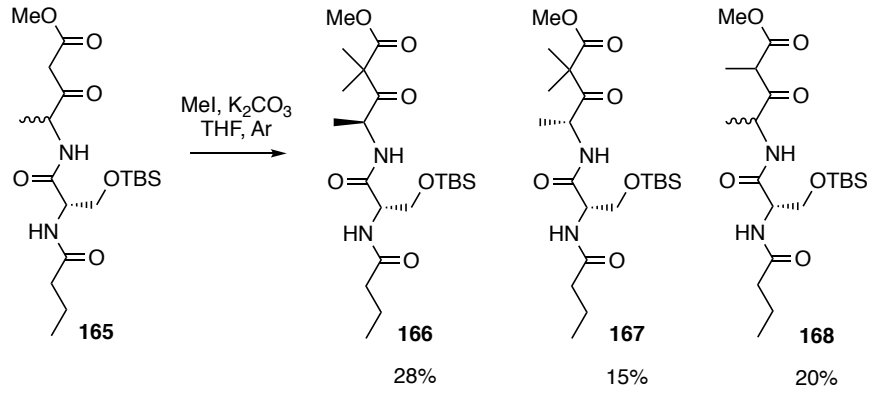
This indicated an issue with the purification of **164** from the lithium hydroxide-mediated ester hydrolysis reaction. It seemed likely that salts which are semi-soluble in organic solvents had carried through as impurities and were inhibiting the coupling reaction. Changing to an acidic workup, where after removal of THF from the reaction mixture the solution was acidified and **164** was extracted from the aqueous layer, increased the yield of the hydrolysis reaction to 82% and no TBS deprotection was observed (scheme 4.9). Subsequent formation of the β-keto methyl ester, using the conditions that had proved most successful in earlier work (1,4-dioxane as solvent, and 50 °C), afforded **165** in 79% yield (scheme 4.9).



Scheme 4.9: Optimised procedure for the synthesis of 164 and 165.

4.2.4 Synthesis of 166

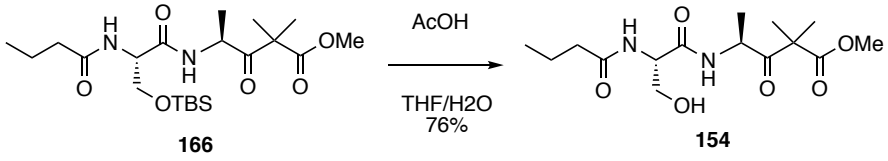
Dimethylation to form the α-dimethyl β-keto methyl ester **166** was attempted using iodomethane and potassium carbonate, following the precedent set during the synthesis of **157** and **144**. The highest yield was obtained when the reaction was conducted at room temperature for 43 hours. Elevating the reaction temperature to 40 °C or increasing the reaction time further caused increased byproduct formation, according to TLC analysis. A mixture of diastereomeric mono- and dimethylated products was obtained, due to racemization of the alanine α-stereocentre. Separation of the mixture on a silica gel column yielded 28 and 15%, respectively, of the dimethylated diastereomers **166** and **167**, and 20% of the monomethylated product **168** as an inseparable mixture of diastereomers (scheme 4.10). The ability to separate **166** and **167**, likely due to the presence of the TBS protecting group, offers a distinct advantage over the previous convergent synthesis of **139** where separation of the diastereomers was not feasible using silica gel chromatography. While **167** and **168** may have also proved to be suitable substrate precursor analogues, this was not investigated due to time constraints.



Scheme 4.10: Products isolated from the methylation of 165 with iodomethane and potassium carbonate.

4.2.5 Synthesis of 154

To selectively remove the TBS group without hydrolyzing the methyl ester of **166**, very mild acidic conditions were used, dissolving **166** in a mixture of 3:3:1 acetic acid:water:THF. This yielded the desired α-dimethyl β-keto methyl ester **154** in 76% yield (scheme 4.11).



Scheme 4.11: TBS deprotection of 166 with acetic acid.

The main problem encountered in the development of the linear synthetic route was the methylation reaction, which produced a mixture of diastereomeric mono- and dimethylated products. Doubtless this reaction could be optimized to produce predominantly the dimethylated products **166** and **167**, albeit still as a mixture of diastereomers. However, this was not investigated due to time constraints.

4.2.6 Basic hydrolysis of 154

Barbas and coworkers demonstrated that a stable lithium salt **169** of an α-alkyl-β-keto ester (figure 4.10A) could be formed using 0.95 equivalents of dilute aqueous lithium hydroxide solution (30 mM).[83] These conditions were therefore investigated for the hydrolysis of **154**. Only ions corresponding to the decarboxylation product **171**, and not its progenitor **170**, were observed when aliquots of the reaction mixture were analysed by MS, suggesting that **170** is highly susceptible to decarboxylation (scheme 4.10B).

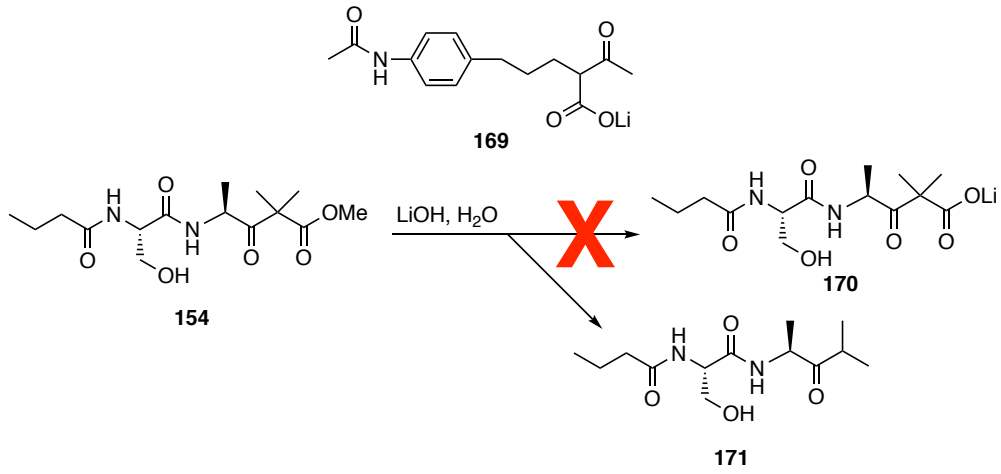


Figure 4.10. A: Structure of a lithium α-alkyl-β-keto carboxylate produced via ester hydrolysis under mild conditions.[83] B: Attempted formation of lithium salt 170 from 154 using analogous conditions. 171 was the only observed product.

The reaction mixture was neutralized with acid, and the product was isolated by silica gel chromatography. ^{1}H and ^{13}C NMR spectroscopic analyses as well as HRMS, confirmed it was **171**. It was thus concluded that base hydrolysis is not an effective strategy for producing EpnF substrates and in all further experiments PLE was used to produce carboxylic acid **172** *in situ*.

4.2.7 Conversion of 154 to an epoxyketone using pig liver esterase and EpnF

UHPLC-ESI-Q-TOF MS analysis of the PLE-catalysed hydrolysis of **154** at pH 8 indicated that the carboxylic acid **172** was produced (figure 4.11A). A small amount of **172** was observed in the negative control employing heat-denatured PLE, presumably resulting from either uncatalysed hydrolysis or residual PLE activity (figure 4.11B).

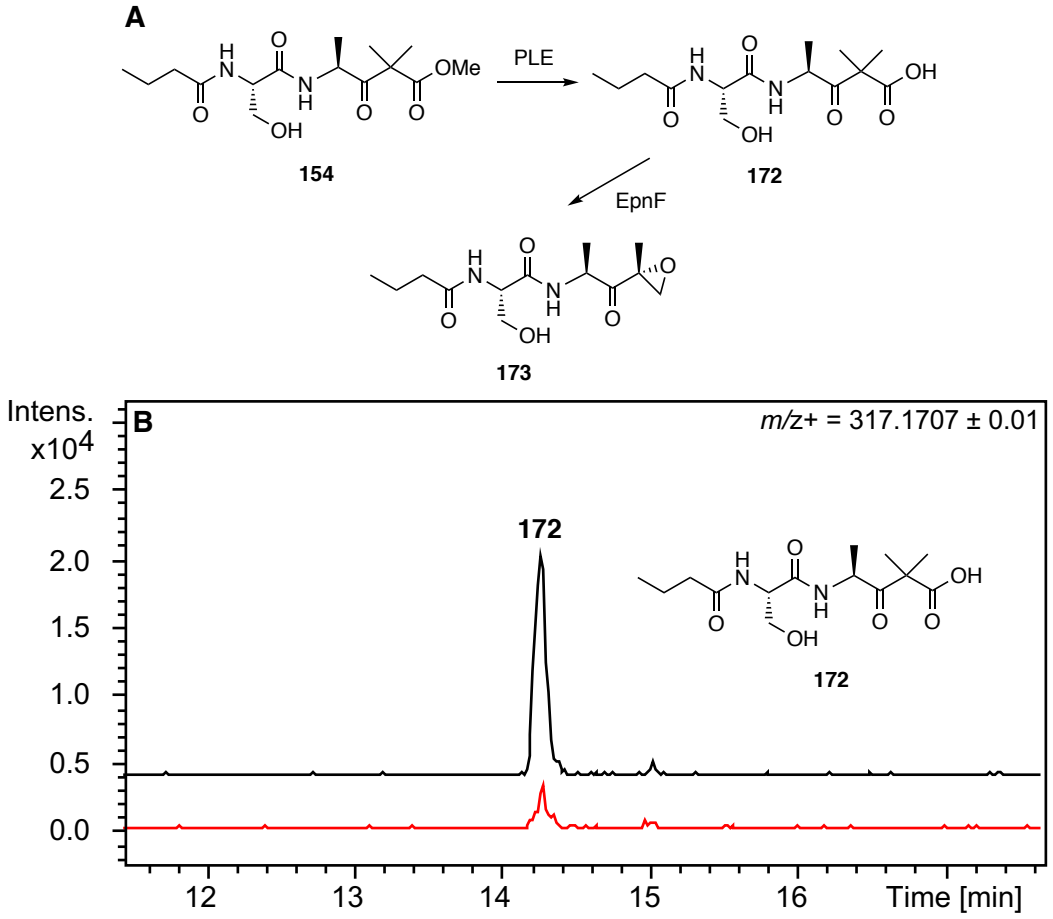


Figure 4.11. A: PLE catalyses the hydrolysis of 154 to produce 172, which is converted to epoxyketone 173 by EpnF. B: EICs at *m/z* = 317.17, corresponding to $[M + H]^+$ for 172, from UHPLC-ESI-Q-TOF MS analyses of assays containing 154 and native (top trace) or heat-denatured (bottom trace) PLE.

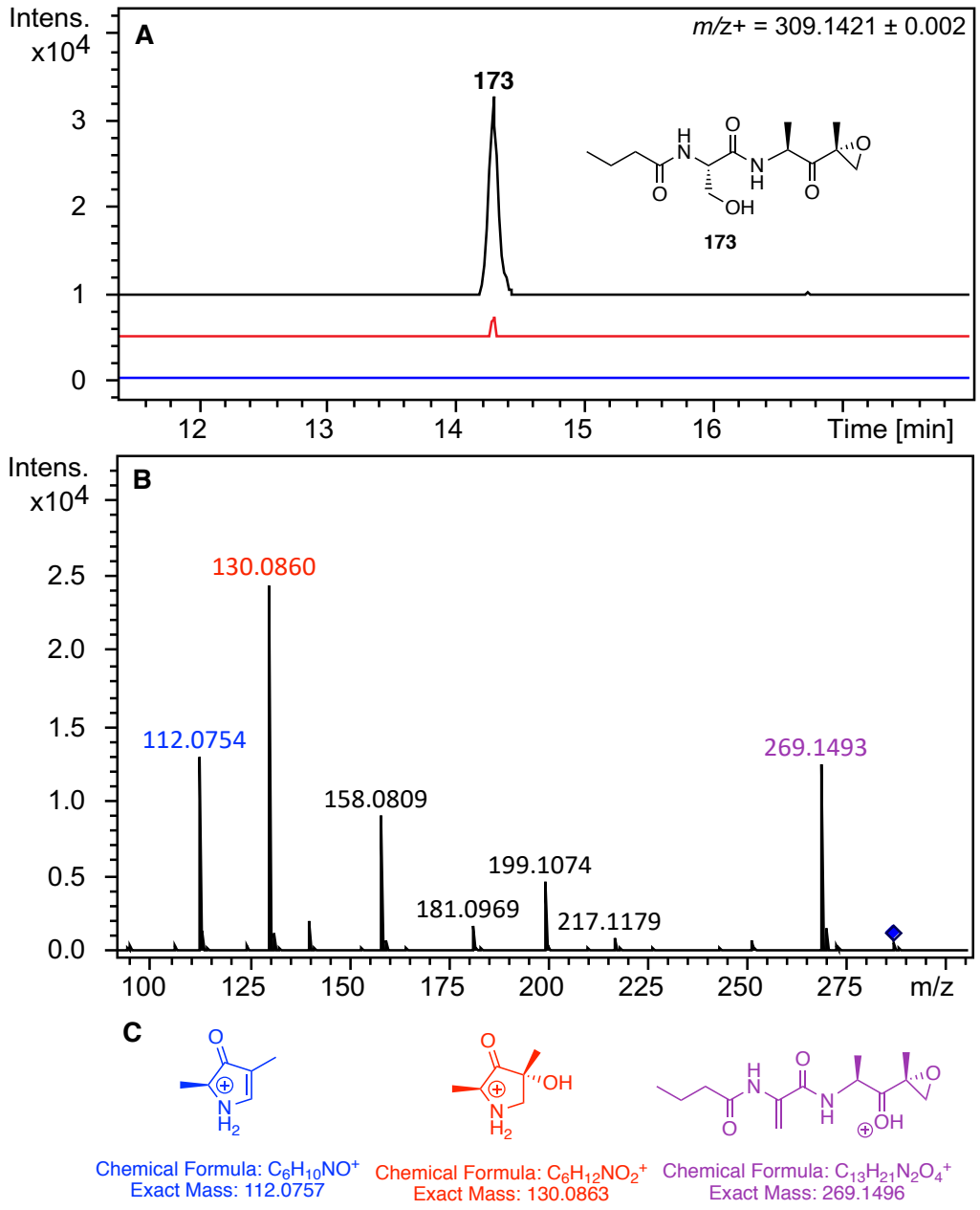


Figure 4.12 A: EICs at *m/z* = 309.14, corresponding to [M + Na]⁺ for 173, from UHPLC-ESI-Q-TOF MS analyses of assays containing 154 and either His₆-EpnF and PLE (top trace), native His₆-EpnF and no PLE (middle trace), or native PLE and heat-denatured His₆-EpnF (bottom trace). B: MS/MS spectrum for the *m/z* = 287.16, corresponding to [M + H]⁺ for 173. C: Proposed structures of observed fragment ions.

UHPLC-ESI-Q-TOF MS analyses of reactions containing PLE, EpnF in pH 8 buffer and **154** showed that a species with a molecular formula corresponding to epoxyketone **173** was produced (figures 4.11A and 4.12A). This product was absent from a negative control containing heat-denatured EpnF, but a small quantity was detectable when PLE was omitted from the reaction, potentially due to residual hydrolysis. Analogous fragment ions to those observed for **66** were observed in MS/MS analyses of the product conducted by Dr. Lijiang Song (figure 4.12B), with similar predicted structures (figure 4.12C). These data strongly imply that **172** is a viable substrate for EpnF, which may possess a broad tolerance with respect to the amino acid-derived side chain adjacent to the α-dimethyl-β-keto carboxylic acid functionality of its substrate.

4.2.7.1 Attempted trapping of alanine-derived epoxyketone with *N*-acetylcysteamine

To further characterise the epoxyketone moiety of **173**, an excess of *N*-acetylcysteamine was added to the reaction containing PLE, EpnF and **154**. However, unlike with **78** (section 3.2.4), neither the epoxyketone **173** nor its *N*-acetylcysteamine adduct were observed by UHPLC-ESI-Q-TOF MS analysis. Levels of **154** when *N*-acetylcysteamine had been added were consistent with a heat-denatured PLE negative control with no *N*-acetylcysteamine present, suggesting that *N*-acetylcysteamine inhibits the function or active site of PLE, with it also possessing an amide bond which cannot be cleaved by the enzyme. With signature MS/MS fragment ions obtained correlating **173** with previous epoxyketone compounds already obtained, nucleophilic derivatization as a means of additional epoxyketone characterisation was not investigated further.

4.2.7.2 Attempted overproduction of TmcH and EpnH thioesterase domains

As an alternative to PLE for the *in situ* hydrolysis of α-dimethyl-β-keto methyl esters the TE domains of TmcH and EpnH were investigated. Although thioesters are the natural substrates of TE domains, it has been reported that they are also able to utilise oxoesters.[91] The DNA corresponding to the TE domains in *epnH* and *tmcH* was therefore cloned into *E. coli* expression vectors, as described for EpnF and TmcI. (sections 2.1 and 2.2.1). Attempts to overproduce the corresponding proteins resulted in insoluble inclusion bodies (figure 4.13). As a consequence, this line of enquiry was not pursued further.

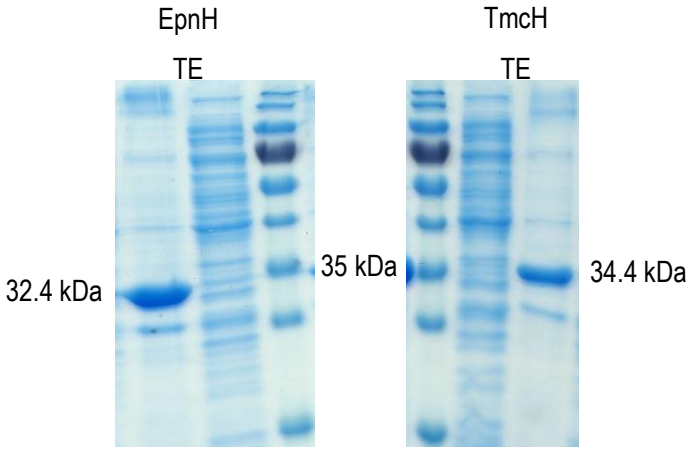


Figure 4.13: SDS-PAGE analysis of the overproduction of the EpnH (34.2 kDa, left hand gel) and TmcH (34.4 kDa, right hand gel) TE domains in *E. coli*. The insoluble protein fractions are in the outside lanes, the soluble protein fractions are in the middle lanes and the molecular weight marker is in the inside lanes.

4.2.7.3 Effect of adding of FAD or FMN to EpnF-catalysed reactions

In the assays of EpnF activity towards **66**, both FAD and FMN had been added to the reaction mixtures to see if there was any effect on the quantity of products formed. However, no difference in the production levels of epoxyketone **78** was observed. In comparison to the vibrant yellow colour of the His_6-EpnF sample used in these assays, the batch of purified enzyme used for the experiments described in section 4.2.7 was pale yellow (figure 4.14), indicating less endogenous flavin cofactor. As efforts to

determine whether the cofactor bound to EpnF was FAD or FMN had proved unsuccessful (section 2.2.3.2), FAD and FMN were added separately to assays containing PLE, EpnF and **154.**

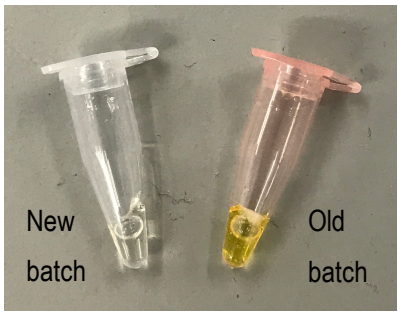


Figure 4.14: A comparison of yellow colour between batches of EpnF, indicating a greater amount of endogenous flavin present in the older purified batch.

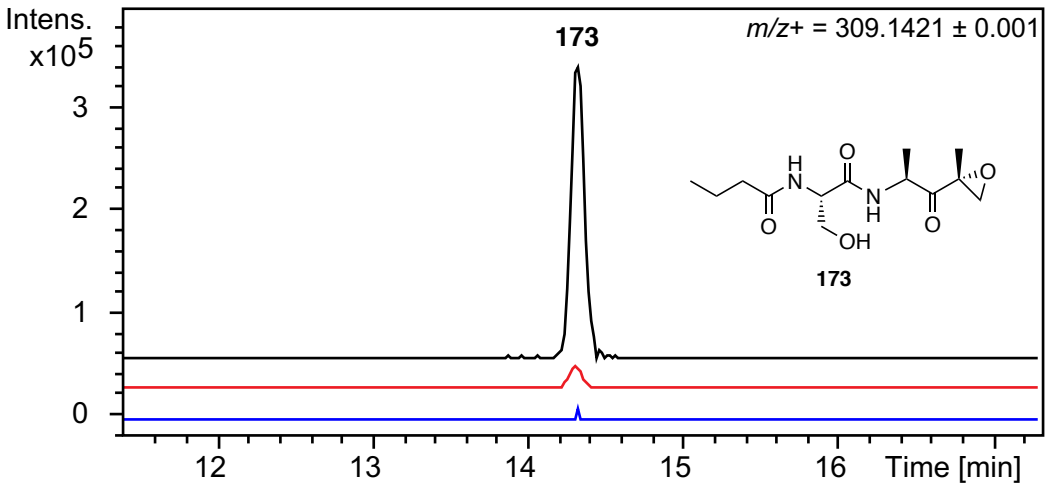


Figure 4.15 A: EICs at *m/z* = 309.14, corresponding to $[M + Na]^+$ for 173, from UHPLC-ESI-Q-TOF MS analyses of assays containing 154, EpnF, PLE and either FAD as an additive (top trace), no additive (middle trace) or FMN as an additive (bottom trace).

UHPLC-ESI-Q-TOF MS analyses revealed a greater quantity of epoxyketone **173** in the assay containing added FAD than the positive control to which no exogenous flavin cofactors had been added (figure 4.15). In contrast, addition of FMN to the

reaction mixture reduced the production of **173** relative to the positive control. This indicates that the flavin cofactor utilized by EpnF is FAD.

4.3 Conclusions

Synthetic routes to two α-dimethyl-β-keto methyl esters, **139** and **154**, were developed. These compounds were converted to β-keto acid substrates (**66** and **172)** for EpnF via *in situ* PLE-mediated ester hydrolysis, overcoming the problem of decarboxylation of the substrate discussed in chapter 3. Although the PLE and EpnF-catalysed reaction cascade is far from optimized, the results obtained indicate that it has the potential to be developed into an efficient catalytic system for production of epoxyketones.

The final peptide coupling step in the convergent synthesis of **139** was found to produce a lactam byproduct. By adopting a linear synthetic route to **154**, this problem was avoided. In addition, the synthesis also allowed disatereomeric separation to give TBS-protected α-dimethyl-β-keto methyl ester **166**. While this may be a result of changing from L-leucine to L-alanine amino acid side chains, it is more likely a result of the TBS-protection itself, which may prove a valuable strategy for similar purification of disastereomers in future analogue syntheses – though leucine methyl ester substrate precursors would have to be prepared through this method to confirm this is the case. α-Dimethyl-β-keto acid **171** produced by *in situ* PLE-catalysed hydrolysis of **154**, was converted to the corresponding epoxyketone by EpnF, showing that the enzyme tolerates alterations to the leucine-derived side chain of **66**. The synthetic procedure developed for **154** could be used in future to produce further substrate analogues, allowing the substrate tolerance of EpnF to be more extensively evaluated.

Chapter 5: Mechanistic investigations of EpnF

5.1 Proposed mechanism of EpnF

In proposing a mechanism for EpnF, several factors need to be taken into consideration. Firstly, the α-dimethyl-β-keto carboxylic acid substrate needs to be converted into an intermediate which could conceivably undergo epoxidation. Since there are no additional cofactors, this process must result in FAD being reduced and converted to a species that could epoxidize a reactive substrate intermediate. The ketone functionality of an α-dimethyl-β-keto carboxylic acid has been described as electron sink, facilitating decarboxylation by subsequent formation of an enolate anion.[54] Dr Zabala's labelled precursor incorporation experiments (figure 1.13) showed that while both of the α-methyl carbon atoms incorporated by the cMT domain of TmcH into TMC-86A, the carbon atom of the terminal carboxyl group is lost.[42] When these facts are combined with the knowledge that **66** is prone to spontaneous decarboxylation to produce **71**, a plausible starting point for the EpnF-catalysed reaction is decarboxylation of an α-dimethyl-β-keto carboxylate to form the corresponding enolate **177** (figure 5.1).

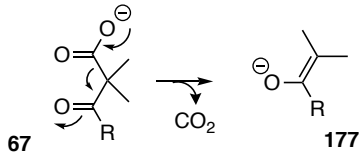


Figure 5.1: Decarboxylation of the β-keto carboxylate substrate of EpnF resulting in formation of enolate 177.

With enolate **177** as a potential intermediate in the EpnF-catalysed reaction, the next question is how can it serve as reductant? It was proposed that the enolate can transfer a hydride ion from one of its β-carbon atoms to enzyme bound FAD. This produces the reduced flavin anion **179** and converts the enolate into the α,β-unsaturated ketone **178** (figure 5.2). Similar chemistry is well-established for flavin-dependent dehydrogenases (section 1.3.2.2).[56]

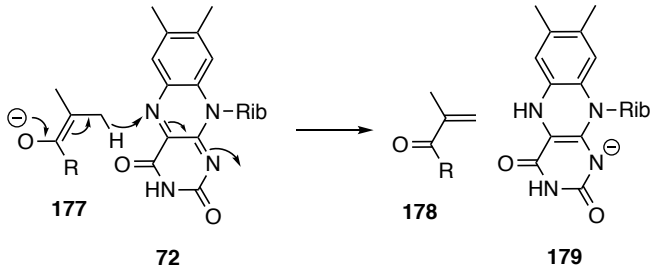


Figure 5.2: Reduction of FAD cofactor by enolate intermediate to give *α,β*-unsaturated ketone and anionic reduced flavin.

179 must be converted to a species capable of epoxidation, such as 4a-hydroperoxyflavin **182** which is widely implicated in a variety of monooxygenation reactions.[92] This reactive moiety forms through initial single electron transfer from the resonance form **180** of anionic reduced flavin to dioxygen to generate flavin-semiquinone **181** and superoxide. Coupling of the radicals in these two species gives anionic 4a-peroxyflavin, which undergoes protonation to give 4a-hydroperoxyflavin **182** (figure 5.3).

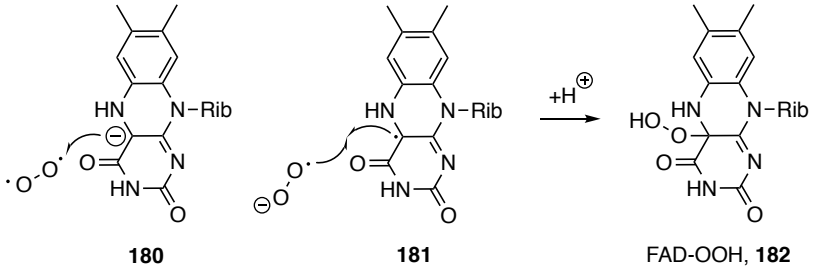


Figure 5.3: Single electron transfer and radical coupling to form 4a-hydroperoxyflavin 182.

Epoxidation of electron-rich alkenes by 4a-hydroperoxyflavin is known, and is proposed to proceed by a peracid-like mechanism.[59,93] With electron-poor alkenes, such as those present in α,β-unsaturated carbonyl compounds, epoxidation is believed to proceed via nucleophilic attack of anionic 4a-peroxyflavin or base-activated 4a-hydroperoxyflavin.[94,95] In the case of **178**, Michael-addition to the α,β-unsaturated ketone would result in the enolate **183**. This could undergo 1,3-elimination to form

the epoxyketone **184** and 4a-hydroxyflavin **185**, which can eliminate water to regenerate the flavin cofactor **72** in its oxidized form.[59] Putting these various processes together, a catalytic cycle for EpnF involving a decarboxylation-dehydrogenation-monooxygenation cascade can be proposed (figure 5.4).[42]

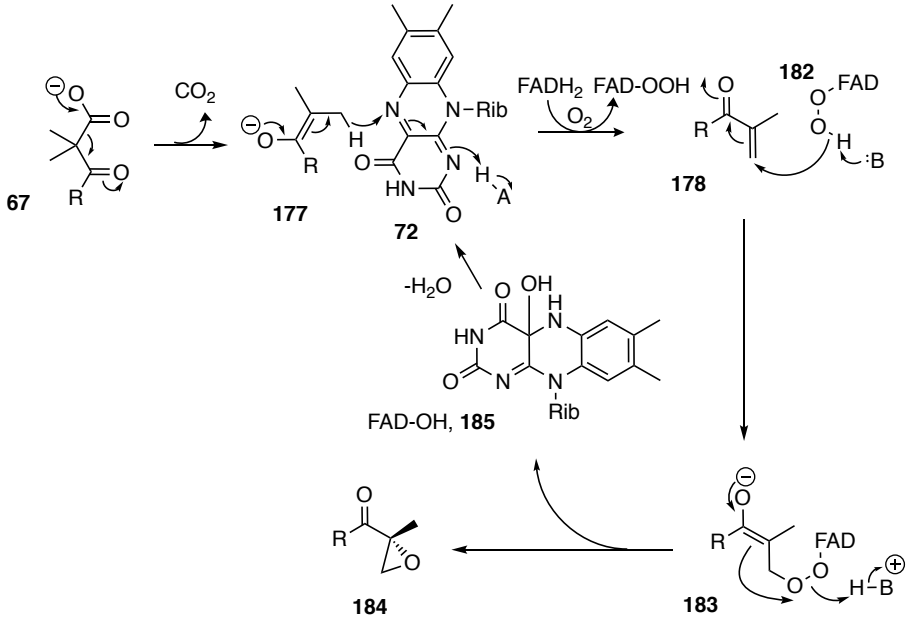


Figure 5.4: Proposed mechanism for conversion of α-dimethyl-β-keto acids 67 to epoxyketones 184 by FAD-dependent enzyme EpnF.[42]

5.1.1 Detection of α,β-unsaturated ketone intermediates in the EpnF catalytic cycle

To obtain evidence for the above mechanistic proposal, it was investigated whether α, β-unsaturated ketone intermediates can be detected in the reactions of **66** with EpnF and **154** with EpnF plus PLE. UHPLC-ESI-Q-TOF MS analyses revealed the presence of compounds with molecular formulae corresponding to **79** and **186** in the transformations involving **66** and **154**, respectively. These α,β-unsaturated ketones were absent from control reactions containing heat-denatured EpnF.

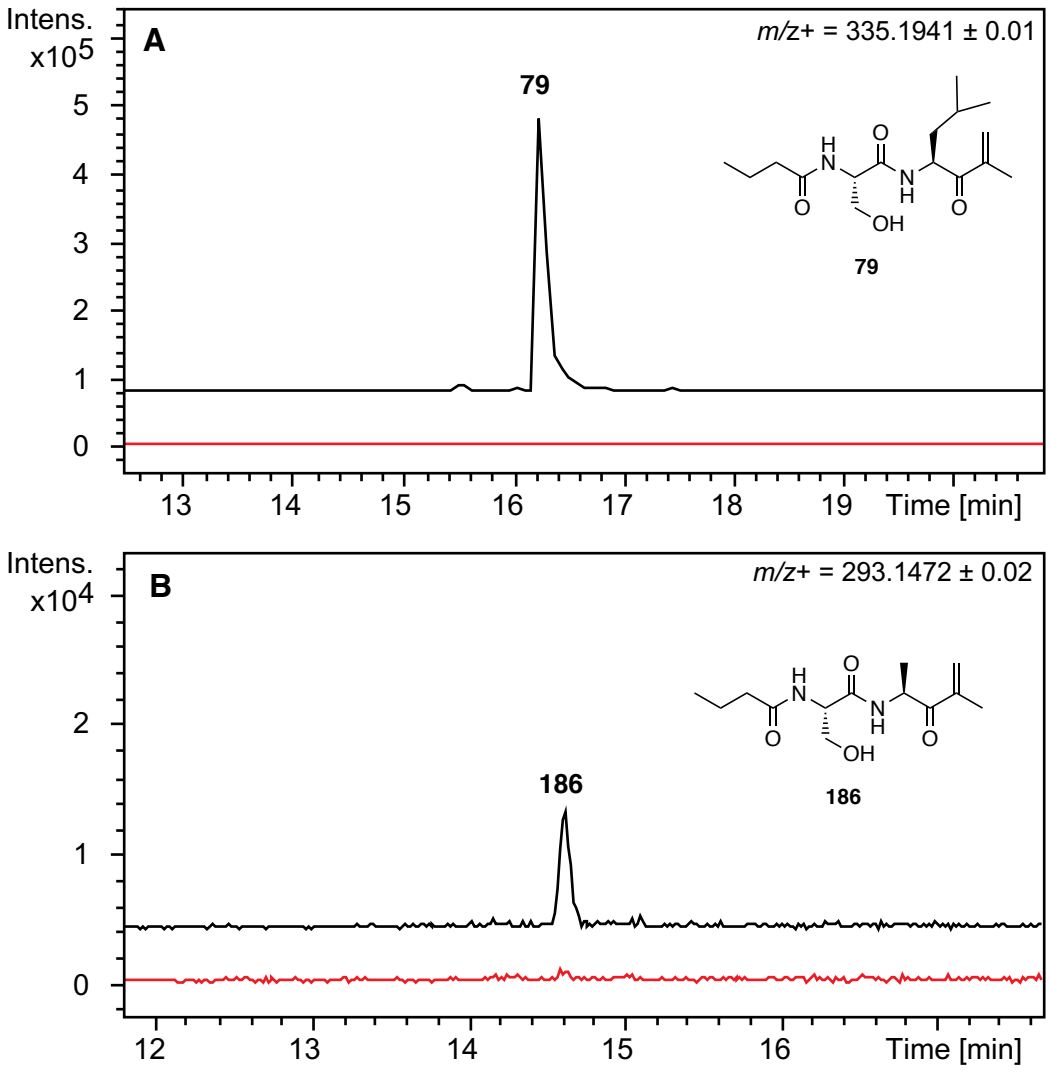
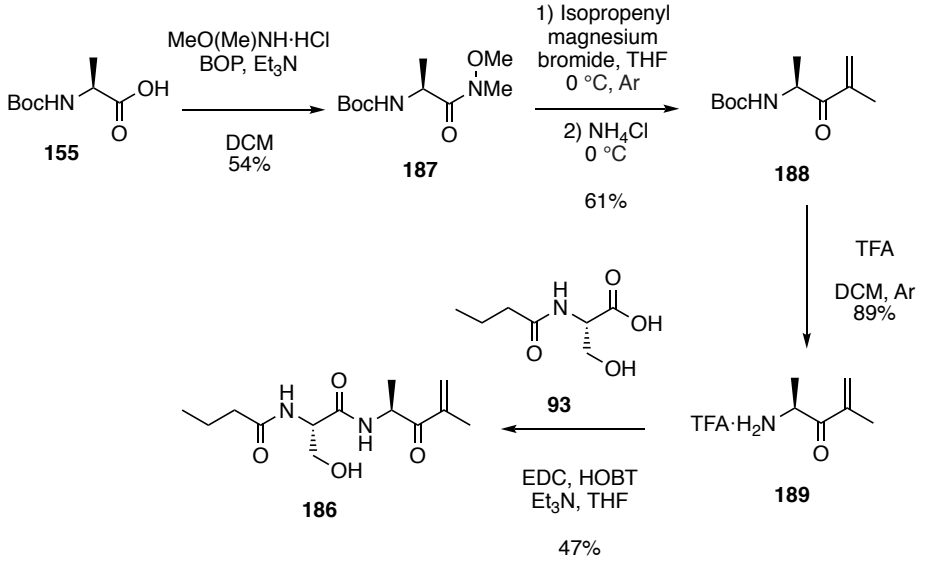


Figure 5.5. A: EICs at *m/z* = 335.18, corresponding to $[M + Na]^+$ for 79, from UHPLC-ESI-Q-TOF MS analyses of assays containing 66 native (top trace) or heat denatured (bottom trace) His_6-EpnF. B: EICs at *m/z* = 293.15, corresponding to $[M + Na]^+$ for 186, from UHPLC-ESI-Q-TOF MS analyses of assays containing 154, PLE and either native (top trace) or heat denatured (bottom trace) His_6-EpnF.

To confirm the structure of **186** an authentic standard was synthesized by building on the synthetic methods developed previously (section 4.2.1.1).

5.2 Synthesis of an authentic standard of the α,β-unsaturated ketone intermediate

Several of the procedures developed in earlier work, such as the Weinreb-Nahm ketone synthesis and the BOC-deprotection (section 3.1.2.3), were directly applicable to the synthesis of **186**.[42,74] The synthesis of **186** was accomplished in 4 steps from *N*-BOC-L-alanine **155** and *N*-butyryl-L-serine **93** in an overall yield of 14% (scheme 5.1).



Scheme 5.1: Synthetic scheme to produce α,β-unsaturated ketone mechanistic probe 186.

5.2.1 Comparison of 186 with the proposed intermediate

The authentic standard of **186** had the same retention time as the putative α,β-unsaturated ketone intermediate detected in reactions containing **154**, PLE and EpnF (figure 5.6), providing strong support for the proposed catalytic mechanism of EpnF.

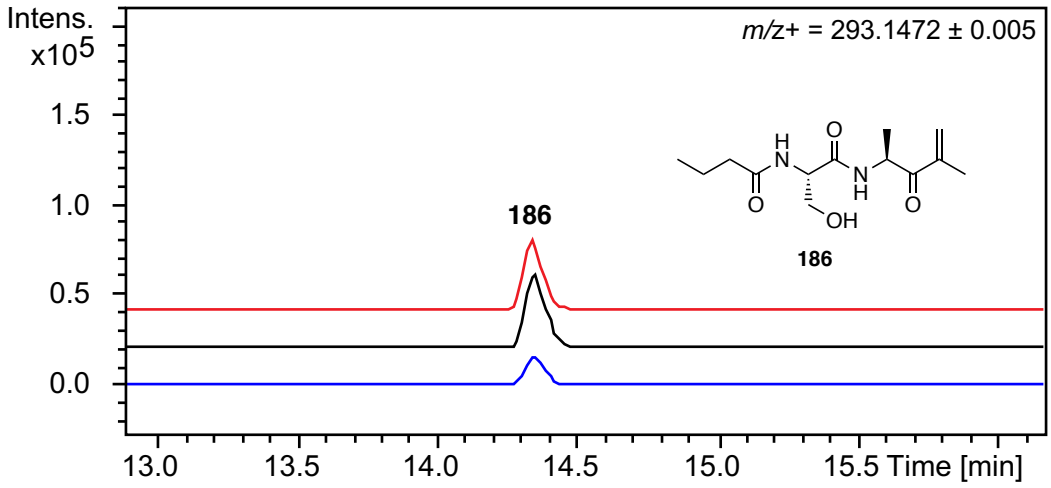


Figure 5.6: EICs at *m/z* = 293.15, corresponding to [M + Na]⁺ for 186, from UHPLC-ESI-Q-TOF MS analyses of a mixture of the authentic standard and the supernatant from the enzymatic reaction (top trace), the authentic standard (middle trace) and the supernatant from the enzymatic reaction (bottom trace).

5.3 Further attempts to probe the EpnF catalytic mechanism

The proposed mechanism for EpnF involves an α,β-unsaturated ketone intermediate that reacts with 4a-hydroperoxyflavin **182** to form an epoxide (figure 5.3). It was reasoned that reduction of the EpnF-bound FAD may lead to the formation of **182**, which could be incubated with **186** to produce epoxyketone **173**. As a consequence, the reduction of EpnF-bound FAD with nicotinamide adenine dinucleotide (NADH) and nicotinamide adenine dinucleotide phosphate (NADPH) was investigated.

5.3.1 Reduction of EpnF-bound FAD with NAD(P)H

NADH and NADPH both have UV absorbance maxima at 340 nm. A decrease in absorbance at 340 nm can therefore be used to monitor the oxidation of NADH or NADPH to NAD^+ or $NADP^+$, respectively, by an electron acceptor such as the flavin cofactor of EpnF (figure 5.7).[96,97]

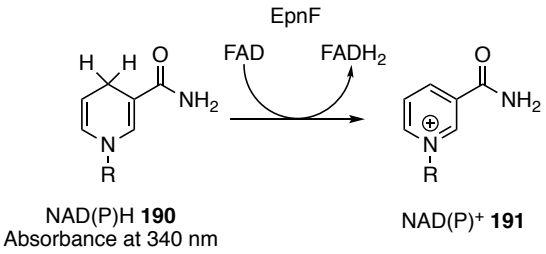


Figure 5.7: Reduction of the EpnF-bound flavin cofactor with NAD(P)H can be monitored by measuring the decrease in absorbance at 340 nm

A steady decrease in absorbance at 340 nm was observed when His_6-EpnF was incubated with NADH, whereas no change in absorbance occurred when the enzyme was incubated with NADPH (figure 5.8). This suggests that EpnF may have evolved from a flavin-dependent monooxygenase, as it appears to have retained the ability to specifically bind NADH and transfer hydride from it to the bound flavin cofactor, even though such as reaction is not required for epoxyketone formation by the enzyme.

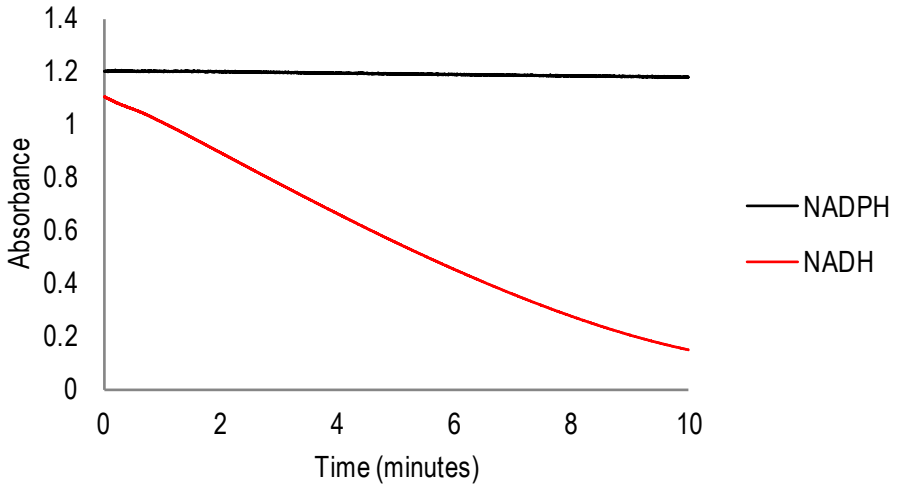


Figure 5.8: Plot of absorbance at 340 nm versus time for incubations of EpnF (40 μM) with NADH (200 μM) and NADPH (200 μM).

5.4 Incubation of EpnF with NADH and 186

In an attempt to obtain further evidence for an α,β-unsaturated ketone intermediate in the EpnF catalytic cycle, the conversion of **186** to **173** by EpnF in the presence of NADH was investigated.

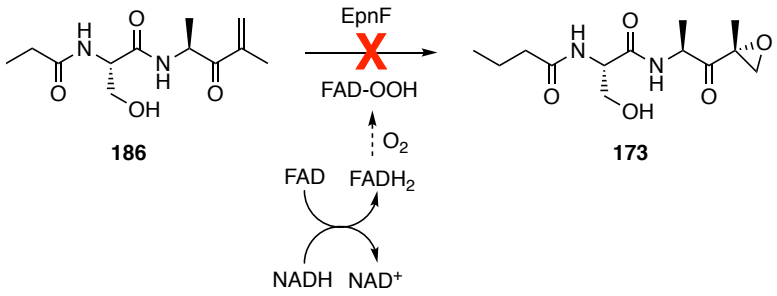


Figure 5.9: Attempted conversion of *α,β*-unsaturated ketone 186 to epoxykketone 173 using EpnF and NADH.

Despite using several different NADH concentrations (0.5/2.0/50.0 mM) no conversion of **186** (1 mM) to **173** was observed (figure 5.9). One possible explanation for this is that once the α,β-unsaturated ketone intermediate is released from the active site of EpnF, the enzyme undergoes a conformational change that prevents the intermediate from re-associating with the active site. Another possible explanation is that NADH/NAD^+ bind to the same site of EpnF as the α,β-unsaturated ketone intermediate and act as competitive inhibitors. Kinetic studies of the effect of NADH/NAD^+ on the PLE and EpnF-catalysed conversion of **154** to **173** could be employed to investigate the latter possibility, whereas structural studies of EpnF may shed light on the former.

5.5 Conclusions

Evidence for an α,β-unsaturated ketone intermediate in the EpnF catalytic cycle was obtained by UHPLC-ESI-Q-TOF MS analyses of previous reactions involving conversion of **66** and **154** to their corresponding epoxyketones. The structure of the

intermediate **186** in the reaction utilising **154** as a substrate was confirmed by chromatograhic comparisons with a chemically synthesised authentic standard.

The flavin cofactor of EpnF could be reduced with NADH but not NADPH, suggesting that it may have evolved from an NADH-dependent monooxygenase. Attempts to convert the α,β-unsaturated ketone intermediate **186** to the corresponding epoxyketone using EpnF in the presence of NADH were unsuccessful. The reason for this is unclear and merits further investigation.

The data obtained are consistent with a catalytic mechanism for EpnF involving decarboxylation of the β-keto acid substrate to form an enolate that transfers hydride from its β-carbon to the flavin cofactor. Reaction of the resulting $FADH_2$ with dioxygen forms a flavin hydroperoxide that epoxidises the carbon-carbon double bond of the α,β-unsaturated ketone resulting from oxidation of the enolate. This likely proceeds via Michael addition of the deprotonated hydroperoxide to the α,β-unsaturated ketone followed by 1,3-elimination.

Chapter 6: Conclusions and future work

6.1 Completion of aims

6.1.1 Enzyme cloning and purification

At the onset of this project, one of the first aims was to clone and purify both the flavin-dependent TmcF and the cytochrome P450 TmcI. While TmcF was overproduced entirely in insoluble exclusion bodies, TmcI by was overproduced in a soluble form as a His_6-fusion. In light of this a homologue from the eponemycin biosynthetic gene cluster EpnF, with 89% amino acid sequence identity to TmcF, was selected for cloning. Later, unpublished research by Dr. C. Huang demonstrated that the eponemycin biosynthetic gene cluster can also produce TMC-86A, validating EpnF's suitability as a homologue (section 1.3.1). EpnF was soluble, and purified in near homogeneity as a His_6-fusion.

The identity of both proteins was confirmed by peptide mass fingerprinting, MS analysis and their elution profile in gel filtration chromatography, the latter to determine their oligomerisation state. UV-vis absorbance revealed the presence of a flavin cofactor in EpnF, with later assays demonstrating this to be FAD rather than FMN. Both MS analysis and UV-vis absorbance revealed the mass and Soret band absorbance of a heme cofactor in the cytochrome P450 TmcI.

Through the methods proposed, EpnF and TmcI can be purified for additional biochemical investigation, with a good degree of characterisation data available to confirm the identity of the purified enzymes.

6.1.2 Synthesis of putative substrates, and *in vitro* assays

Through a peptide coupling approach, both the potential substrates **71** and **66** were successfully synthesised. An issue arising from **66** was its anticipated spontaneous degradation to isopropyl ketone **71** through decarboxylation, facilitated by intramolecular hydrogen bonding in a cyclic transition state in the α-dimethyl-β-keto carboxylic acid (section 1.3.2.1). A new aim arising from this issue was to purify either a stable analogue of the substrate, or a stable substrate precursor.

Investigation of *in vitro* activity of **66** and **71**, with appropriate negative controls, revealed **66** and not **71** to be a putative substrate for EpnF. From this assay, epoxyketone **78** was produced, confirmed by UHPLC-ESI-Q-TOF MS analysis. While a suitable quantity for NMR analysis could not be purified in a scale-up assay using EpnF and **66** due to the aforementioned substrate instability, MS/MS fragmentation and derivatization using *N*-acetylcysteamine provided a measure of characterisation data for epoxyketone **78**. This result corroborated the findings of Wenjun and coworkers, who had determined through heterologous expression in *E. coli* that EpnF was vital for epoxyketone formation.[79]

As TMC-86A and eponemycin natural products are produced as a single diastereomer, this is likely the case for the epoxyketone **78**, supported by the elution of species with the correct molecular formula in a single peak in UHPLC-ESI-Q-TOF MS analysis of the assay supernatant. From an industrial perspective, chemoenzymatic synthesis using this class of flavin dependent enzymes could ensure a high degree of stereoselectivity in the epoxyketone forming reaction. By comparison only 65% of the desired enantiomer is produced via industrial synthesis (scheme 1.1).

As the role of TmcI had yet to be determined, it was used in an assay with **66**, EpnF and an appropriate ferredoxin/ferredoxin reductase redox system from spinach. UHPLC-ESI-Q-TOF MS analysis revealed that TmcI was responsible for hydroxylation of the α-methyl group of the TMC-86A epoxyketone pharmacophore, producing **42**. This result was confirmed by MS/MS analysis, which when compared to the MS/MS fragmentation of TMC-86A and **78** produced a signature pattern of fragment ions.[42]

It is worth noting that while the epoxomicin biosynthetic gene cluster from *Goodfellowiella coeruleoviolacea* contains a homologue to both TmcI and EpnI in the form of EpxC, no similar hydroxylation of the α-methyl group of the epoxomicin epoxyketone is observed (table 1.1).[41] This indicates that EpxC is likely a mutated or inactive variation of TmcI/EpnI.

6.1.3 Synthesis of a stable substrate precursor analogue

With the issue of decarboxylative degradation presenting itself in the synthesis of **66**, a new aim arising for the project was to synthesise a stable substrate precursor, which was attempted for methyl ester **139**. Initially a Meldrum's acid derived procedure was chosen for the two carbon homologation step to produce β-keto methyl ester **143**, however yield for this compound was significantly improved with a CDI mediated methyl potassium coupling. **139** was obtained in a total yield of 4.9%, with attempts at optimization of the last coupling step yielding the lactam byproduct **151**, likely due to the Thorpe-Ingold effect leading to an increase in the intramolecular reaction between the amino group and β-keto methyl ester of **145**. This would need to be addressed in future syntheses of these methyl ester substrate precursors.

In order to convert **139** into the substrate **66**, PLE was investigated as a mild esterase. Turnover to **66** was observed with PLE, and as purification of a more specific esterase in the form of the TE domains of TmcH and EpnH was unsuccessful, it was selected as the esterase of choice for an attempted one pot epoxyketone synthesis with PLE.

In a paired assay of PLE and EpnF, **139** was successfully converted into a compound with the mass of epoxyketone **78**, which was absent in controls with heat-denatured EpnF by UHPLC-ESI-Q-TOF MS. As PLE is routinely used in the chiral resolution of esters adjacent to a racemic stereocenter, there is however the possibility that PLE was selectively hydrolysing the *SR*-disatereomer of **139** rather than the putatively favoured *SS*-diastereomer.[82] This merits further investigation into the possibility of whether EpnF accepts the *SR*-disastereomer, either by purification of the separate diastereomers or by identification of an esterase lacking a bias towards a specific stereochemistry.

While this process was not efficient, it demonstrated a proof of concept of a one pot synthesis of epoxyketone compounds from α-dimethyl-β-keto methyl esters. Optimization of this one pot reaction is also a natural focus for future investigation, and could lead to a scalable method of epoxyketone production for industrial, due to the inherent improved stability of the methyl ester precursors as opposed to the putative natural α-dimethyl-β-keto carboxylic acid substrates.

6.1.4 Investigation into substrate specificity of EpnF

To investigate the substrate specificity of EpnF, an analogue based on alanine **154** was selected as the synthetic target, due to a natural epoxyketone clarepoxcin A also having alanine in this position (table 1.1).

While initially a convergent synthesis was attempted at the same time as **139**, coupling by peptide synthesis two halves of the desired compound, lactam byproduct formation was also observed for the alanine synthetic route producing **159**. In a briefly attempted linear synthesis of **139**, presence of the hydroxyl group of an attached serine residue appeared to inhibit the reaction of the leucine carboxylic acid with CDI and methyl potassium malonate, suggesting protection of this hydroxyl group was necessary in other attempts at analogue synthesis. Therefore a linear synthetic route was proposed to **154**, using TBS protection of the serine side chain. Diastereomeric separation in the subsequent α-dimethylation reaction of the β-keto methyl ester **165** was accomplished through, likely arising from the presence of the sterically large non-polar TBS group. This offered a significant advantage of the previously proposed route to **139**, where separation of diastereomers could not be accomplished through silica gel chromatography. This linear synthesis proved successful, providing an attractive synthetic route to future analogues.

154 was in the presence of PLE demonstrated turnover to carboxylic acid **171**. In a one pot reaction with both PLE and EpnF, the mass of a new epoxyketone compound **172** was observed by UHPLC-ESI-Q-TOF MS analysis of the assay supernatant. This demonstrates that EpnF has a broader substrate tolerance than the leucine amino acid residue incorporated in **139** and **66**, and invites more investigation into substrate tolerance of the enzyme.

6.1.5 Authentic standard and mechanistic intermediate

A decarboxylation-dehydrogenation-monooxygenation mechanism for EpnF was proposed (figure 5.4) that resolved the question of how the innate FAD cofactor of EpnF could be reduced in the absence of an additional external cofactor. A substrate

intermediate on this proposed mechanistic pathway, before the reaction with hydroperoxyflavin to produce an epoxyketone, was a reactive α,β-unsaturated ketone species. Analysis of assay supernatant for the EpnF-catalysed reactions with **66**, **139** or **154** revealed the presence of compounds with masses consistent to these reactive intermediates in each instance.

In light of this it was desirable to produce an authentic standard of one of these intermediates, and as syntheses of alanine containing analogues had provided higher yields than their leucine counterparts, synthesis of the α,β-unsaturated ketone **186** was devised. This synthesis was readily completed in total yield of 14%, using procedures established during the course of this project.

Co-injection of **186** with the supernatant of an assay containing PLE, EpnF and **154** and subsequent UHPLC-ESI-Q-TOF MS analysis revealed **186** to be an authentic standard for the corresponding proposed α,β-unsaturated ketone intermediate. This provides significant evidence that the putative decarboxylation-dehydrogenation-monooxygenation mechanism for this class of flavin-dependent proceeds through such intermediates to produce epoxyketone compounds. Briefly, use of **186** as a substrate with EpnF was attempted, which had its flavin cofactor reduced by NADH. However, no turnover was observed in this instance, indicating either that these intermediates cannot re-associate with the active site, or that NADH may inhibit the active site. The latter of these two possibilities may be a result of EpnF evolving from flavin dependent enzymes which require NADH reduction, with NADH still successfully being able to reduce EpnF.

6.2 Future work

6.2.1 Further investigation into substrate tolerance

With a consistent route to substrate synthesis now identified for the methyl ester substrate precursor **154**, and a reproducible one pot reaction using PLE and EpnF established, it would be desirable to further evaluate the substrate tolerance of EpnF to produce novel analogues of epoxyketones. These epoxyketone compounds may have improved efficacies over existing clinically used pharmaceuticals such as

carfilzomib (figure 1.7). Possible areas for functional group variation of substrates for EpnF are the amino acid side chain adjacent to the α-dimethyl-β-keto methyl ester functionality, different α-alkyl substitution between the carbonyl functionalities, and use of a different peptide backbone, as demonstrated in structure **192** (figure 6.1). Two particularly interesting substrates to produce may be analogues with either an L-tryptophan or an L-phenylalanine residue adjacent to the α-dimethyl-β-keto methyl ester, compounds **193** and **194**. A tryptophan side chain is observed in the natural product tryptopeptin A (table 1.1), while an L-phenylalanine side chain can be observed in the second generation epoxyketone pharmaceutical oprozomib.[23,98] Additionally, chemoenzymatic total synthesis of Carfilzomib would further evidence the industrial relevance of this class of enzymes.

Lastly, as previously mentioned the Histone Deacetylase HDAC8 inhibitor trapoxin A also has an epoxyketone pharmacophore at the end of a long alkyl chain (figure 1.4B). It would be interesting to see if a similar α-methyl-β-keto methyl ester functionality at the end of a long alkyl chain such as **195** could be used in the one pot PLE and EpnF reaction to produce an epoxyketone, as this would make EpnF capable of yielding another anticancer epoxyketone compound with a different biological target (figure 6.1).[29]

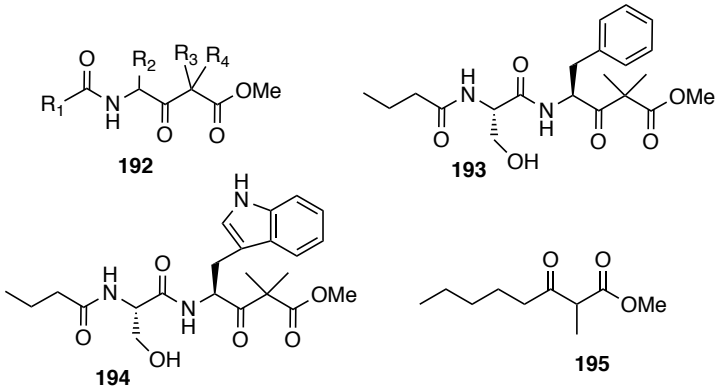


Figure 6.1: Potential substrate candidates to explore EpnF substrate tolerance.

6.2.2 Identification of an esterase with increased activity

Acquiring a range of esterases and screening them for hydrolytic activity against an α-dimethyl-β-keto ester substrate analogue would help to identify an ideal esterase for pairing in a one pot reaction with EpnF. One means of high-throughput screening of esterase activity is achieved by use an ester which fluoresces upon hydrolysis. An example of fluorescent tag is an umbelliferone based cyanohydrin ester, which degrades to umbelliferone **200** upon hydrolysis and can be monitored by measuring absorbance at 440 nm (figure 6.2).[99] Synthesis of an appropriate analogue, such as **196**, would allow for screening of esterases using this procedure.

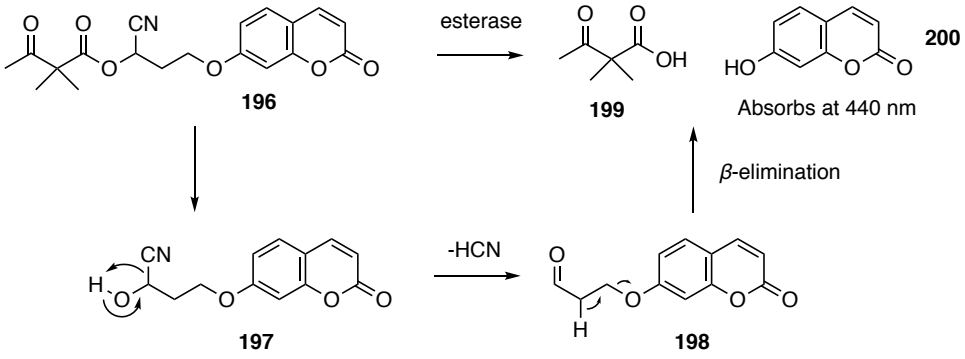


Figure 6.2: An example of high-throughput esterase screening by use of a fluorescent tag.

6.2.3 Crystal structure of EpnF

With the seemingly novel mechanism of EpnF, it would be very interesting to evaluate the nature of the binding pocket within the active site which facilitates this process. With the synthesis of a variety of ligands which may aid crystal formation already completed (**139**, **154**, **186**), and NADH potentially having serving as an active site inhibitor (section 5.4), formation of crystals for X-ray crystallographic study would be highly desirable. With a crystal structure produced, techniques like a directed evolution approach could be employed to increase the tolerance of EpnF for unnatural substrates.[100]

Chapter 7: Experimental procedures

7.1 General

7.1.1 Chemical reagents

Chemical reagents were bought from Sigma Aldrich, Fisher Scientific, Acros Organics, Alfa Aesar or Chem-Impex international. Solvents were sourced from Sigma Aldrich or VWR Chemicals. (*S*)-1-(Benzyloxy)-2,2,6-trimethyl-1,3-dioxoheptan-4-ammonium trifluoroacetate **122** was synthesized by Dr. D. Roberts.[42] HR-MS, 500 MHz ^{1}H-NMR and 125 MHz ^{13}C-NMR data were obtained by the departmental MS and NMR services, run by Dr. Lijiang Song and Dr. Ivan Prokes respectively.

7.1.2 Instrumentation

UHPLC-MS and UHPLC-MS/MS analyses were carried out using a Bruker MaXis Impact ESI-TOF-MS connected to a Dionex 3000 RS UHPLC instrument fitted with an Agilent Zorbax Eclipse Plus C18 column (100 × 2.1 mm, 1.8 μm, 25 °C) for chemical compounds and Chromatography Technologies ACE 3 C4-300 reverse phase column (100 × 2.1 mm, 1.8 μm, 25 °C) for purified proteins.[42]

NMR spectra of synthetic compounds were measured on Bruker AV 300, 400 or 500 MHz spectrometers. Low resolution mass spectra of synthetic compounds were recorded using an Agilent 6130B single Quad spectrometer. High resolution mass spectra of synthetic compounds were measured on a Bruker MaXis ESI-Q-TOF-MS spectrometer.[42] Infrared spectra were recorded using a PerkinElmer Universal ATR Sampling Accessory. Optical rotations were recorded using a Optical Activity AA-1000 Polarimeter.

Centrifugation was conducted using eppendorf 5804 R and 5415 R centrifuges and a Sorvall RC 6 Plus centrifuge fitted with either a SLA-3000 or SS-34 rotor. PCR reactions were conducted using an eppendorf Mastercycler gradient machine. The pH of buffer solutions was measured using a Fisher Hydrus 300 pH meter. The shaker-incubator used during cell growth was a New Brunswick Scientific innova44

incubator. Optical density measurements to determine cell growth were measured with Thermo BioMate3 spectrophotometer.

PCR reactions were conducted using an eppendorf Mastercycler gradient machine. Agarose gel electrophoresis was carried out using a BioRad Wide Mini-Sub Cell GT tank connected to a BioRad Power Pac 300 power supply. Agarose gels resulting from electrophoresis were visualized using under UV light using a UVP imaging system UV transilluminator. Both protein and DNA concentrations were measured using a Thermo Scientific NanoDrop Lite spectrophotometer. Cells were lysed under pressure using a Constant Systems cell disruptor. SDS-PAGE analysis was conducted with a BioRad Mini PROTEAN Tetra Cell tank connected to a BioRad Power Pac 300 power supply.

A Varian Cary 50 Bio UV-Visible spectrophotometer was used for protein UV-visible measurements. Gel filtration chromatography was carried out using an Amersham Biosciences ÄKTA purifier connected to a GE Healthcare HiLoad 16/600 Superdex 200 pg column.[42]

7.1.3 Plasmids, strains and commercial enzymes

Genomic DNA was prepared from *S. hygroscopicus* ATCC 53709 and *S. chromofuscus* ATCC 49982. Competent cells used were either Invitrogen *E. coli* One Shot TOP10 or *E. coli* One Shot BL21 Star (DE3) Chemically Competent Cells. The plasmid vector used was pET151 from an Invitrogen Champion pET151 Directional TOPO Expression Kit. Unless otherwise stated commercial enzymes were purchased from Sigma Aldrich.

7.1.4 Biological reagents and media

Fisher Luria-Bertani (LB) miller media was made to a concentration of 25 g/L in deionized water. Solid LB-agar media was made using 25 g/L LB media and 15 g/L of Becton, Dickinson and Company bacto agar in deionized water. Stock solutions of ampicillin were made to a concentration of 100 mg/mL and sterilized by syringe filtration, and when used to supplement media it was added such that the final

concentration of antibiotic would be 100 μg/mL. Fluka tryptic soy broth (TSB) media for growth of *Streptomyces* species was prepared to the supplier's specifications in deionized water. Biological reagents were purchased from Sigma Aldrich unless otherwise stated.

7.2 Biological procedures

7.2.1 Isolation of genomic DNA

TSB media (50 mL) was inoculated with a *Streptomyces* 10^6 spore stock solution (50 μL) and grown at 30 °C and 180 rpm for 2 days. The culture was centrifuged at 4000 rpm for 10 min and the supernatant discarded.

The cell pellet was washed twice with STE buffer (5 mL - 100 mM sodium chloride, 100 mM Tris pH 8.0, 1 mM ethylenediaminetetraacetic acid – EDTA). Cells were resuspended in both STE buffer (5 mL) and lysozyme (5 mL, 10 mg/mL in STE buffer). Solution was incubated at room temperature for 30 min with gentle shaking. Sodium dodecyl sulfate solution (0.5 mL, 10% *w/v*) was added and the solution was incubated for 5 min with gentle shaking at room temperature. The solution was heated to 70 °C for 5 min and then cooled on ice for 5 min. Potassium acetate (1.25 mL, 5 M) was added and the solution was incubated for 5 min on ice.

Phenol-chloroform-isoamyl alcohol (25:24:1, 5 mL) was added and the solution mixed gently by inversion for 5 min. The solution was centrifuged at 7000 rpm for 10 min at 20 °C. Sodium acetate (1 mL, 3 M) and ice-cold isopropanol (6 mL) were added to the supernatant and mixed by inversion to precipitate the genomic DNA. The solution was centrifuged at 7000 rpm for 10 min at 4 °C, then the supernatant discarded. The pellet was washed with 70% ethanol (100 μL) in deionized water, and centrifuged at 7000 rpm for 10 min at 4 °C discarding the supernatant and leaving the pellet to dry. Once dry the DNA was resuspended with Geneaid RNase in water (500 μL, 10 mg/mL). The concentration was measured using a NanoDrop spectrophotometer and the genomic DNA was stored at -20 °C.

7.2.2 Polymerase chain reactions and agarose gel electrophoresis

Complementary primers (Sigma Aldrich) were ordered for the amplification of *epnF, tmcF* and *tmcI* from the genomic DNA of *S. hygroscopicus* and *S. chromofuscus* respectively (table 7.1), with forward primers utilizing a CACC tag for cloning into a pET151 vector.

Table 7.1: Primers used for amplification (-f or -r) or sequencing (-seq) of *epnF* or *tmcF* and *tmcI* from *S. hygroscopicus* and *S. chromofuscus* respectively.

Primers	Nucleotide sequence (5' → 3')
epnF-f	CACCGTGAGTGACAGCAAATCGG
epnF-r	TCATGCTTCCCCGGGTGAG
epnF-seq	ATCCAGACAGCCGTCGGCTC
tmcF-f	CACCGTGGTGGAACGCCTGGCGGATA
tmcF-r	TCATCGCTCCCCCGGGTGGGTG
tmcI-f	CACCATGGTGACGATCGACCCGAACA
tmcI-r	CTGTCAGACATGGGCGTTCTCT

PCR reactions were prepared using a Roche Diagnostics Expand High Fidelity PCR System kit, using forward and reverse primers for gene of interest (table 7.1) and genomic DNA (Section. 7.2.1.), in the volumes specified in the kit's protocol (table. 7.2).

Table 7.2: Quantities of reagents used for PCR reactions.

Reagent	Volume (µL)
Genomic DNA (100 ng/µL)	2
Buffer 2 ($MgCl_2$)	5
dNTPs (10 mM)	1
Forward primer (10 µM)	1.5
Reverse primer (10 µM)	1.5

High Fidelity polymerase	0.5
DMSO	5
Deionized water	33.5
Total volume	50

Annealing temperatures between 55 °C and 65 °C were screened for all PCR reactions, with *tmcI* and *epnF* favoring 55 °C and *tmcF* favoring 60 °C. The following temperature scheme was used for the PCR reactions, with 35 cycles of steps 2-4, using an eppendorf Mastercycler gradient machine:

1) Denaturation: 94 °C for 5 min
2) Denaturation: 94 °C for 1 min
3) Annealing: 55 °C (*tmcI*/*epnF*) or 60 °C (tmcF) for 1 min
4) Extension: 72 °C for 2 min
5) Final extension: 72 °C for 10 min
6) Stop reaction: 4 °C until ready for use

PCR products were visualized and purified using agarose gel electrophoresis. Agarose gels were prepared by dissolving Appleton D-1 LE agarose (1.0 g) in TBE buffer (100 mL - 4.46 M Tris, 4.45 M boric acid, 200 M EDTA) by microwave heating. The solution was allowed to cool to ca. 50 °C and Biotium GelRed Nucleic Acid Gel Stain (5 µL, ×10000 concentrated) was added. The gel was allowed to set in a gel tray with a comb to ensure well formation.

PCR products were mixed with Thermo Scientific ×6 loading dye (2 µL) and added to separate wells of the agarose gel. The gel was submerged in TBE buffer in a BioRad Wide Mini-Sub Cell GT tank, and ran at 90 V at room temperature for ca. 1 hour. The resulting gel was visualized under UV light, the PCR products excised and purified as per the manufacturer's protocol using a Thermo Scientific GeneJET Gel Extraction Kit. Concentration of the PCR product was measure using a Thermo Scientific NanoDrop Lite spectrophotometer.

7.2.3 Gene cloning, competent cell transformation and plasmidic DNA extraction

PCR product (ca. 2 ng) was added to salt solution (1 μL, 1.2 M NaCl, 0.06 M $MgCl_2$) and pET151 vector (1 μL, 15-20 ng/μL) from an Invitrogen Champion pET151 Directional TOPO Expression kit, making up to a final volume (6 μL) with deionized water. The reaction was left for 30 min at room temperature, and added to an Eppendorf tube of Invitrogen *E. coli* One Shot TOP10 chemically competent cells, stirring with a pipette tip. Cells were left for 5 min on ice before transforming the cells with the plasmid using heat shock (42 °C for 30 sec). The vial was transferred back to ice and LB media (500 μL) was added to the Eppendorf tube. The mini-culture was shaken at 180 rpm and 37 °C for 1 hr. The solution was spread on ampicillin supplemented LB-agar plates and which were then incubated overnight at 37 °C.

From the LB-agar plate five single colonies were picked and added to separate tubes LB media (10 mL) supplemented with ampicillin stock solution (10 μL). The cultures were allowed to grow at 180 rpm and 37 °C overnight.

Glycerol stocks were made by centrifuging this cell culture (1.5 mL) at 13200 rpm for 1 min, discarding the supernatant then re-suspending the cell pellet in LB (500 μL) and glycerol (500 μL, 50% *v/v* in deionized water). Glycerol stocks were stored at -80 °C. Plasmids were extracted from the remaining culture liquid using a ThermoScientific GeneJET Plasmid Miniprep Kit as per the manufacturer's protocol. Concentration was measured using a Thermo Scientific NanoDrop Lite spectrophotometer.

7.2.4 Construct restriction digests and gene sequencing

The restriction enzymes *Pvu*I or *Sac*I were selected to screen potential *epnF* and *tmcI* or *tmcF* pET151 constructs respectively, cleaving both the gene of interest and the plasmid vector once in order to confirm orientation of the gene. Digests were prepared by combing *Pvu*I or *Sac*I (1 μL, 10 units/μL, New England Biolabs), NE 3.1 or 1.1

buffer (1 μL, ×10 concentrated), deionized water (3 μL) and the plasmid construct (5 μL, concentrations between 10 ng/μL and 30 ng/μL) and incubated at 37 °C. Agarose gel electrophoresis (section 7.2.2) was conducted on the digests and they were visualized under UV light. Constructs which appeared to correspond to successful digestion were sequenced by GATC Biotech, using the primers T7 and pET-RP which flank the cloned gene in the construct. The *epnF*-pET151 construct required an additional sequencing primer to sequence the middle section of the gene (*epnF*-seq, table 7.1). The genetic sequence of the constructs was checked for mutations using the ClustalW2 alignment program software, comparing to the original sequence of each gene.

7.2.5 Protein overexpression

Both *epnF*-pET151 and *tmcI*-pET151 plasmid constructs were used to transform Invitrogen *E. coli* One Shot BL21 Star (DE3) chemically competent cells (section 7.2.3). Single colonies from ampicillin supplemented LB-agar plates were selected and transferred to ampicillin supplemented LB media (10 mL), growing overnight at 37 °C and 180 rpm.

The overnight culture was used to inoculate LB media (1.0 L) supplemented with ampicillin stock solution, and cells were grown at 37 °C and 180 rpm until they reached an A_{600} optical density of between 0.5-0.7. Syringe filter sterilized IPTG was added to the culture (final concentration of 500 mM for TmcI, 100 mM for EpnF), in addition to riboflavin (500 mM final concentration) in the EpnF culture and 5-aminolevulinic acid (250 mM final concentration) for the TmcI culture in order to assist formation of the flavin and heme cofactors respectively. The culture was shaken overnight at 180 rpm and 15 °C.

Cultures were centrifuged at 4000 rpm and 4 °C for 30 minutes and the supernatant discarded. The cell pellet was re-suspended in washing buffer (20 mM Tris-HCl, 100 mM NaCl, 10 % glycerol, 20 mM imidazole, pH 8), and then either stored at -80 °C or used in protein purification.

7.2.6 Protein purification by nickel ion affinity chromatography and SDS-PAGE analysis

Ice cold phenylmethylsulfonyl fluoride (PMSF – 1 mM final concentration) was added to the suspended cell pellet (section 7.2.5), and the cells lysed by pressure using a Constant Systems cell disruptor. Cellular debris was pelleted by centrifugation (2 × 20 min, 15000 rpm, 4 °C) and the supernatant passed through a 0.45 μm syringe filter. Filtrate was loaded onto a GE Healthcare HisTrap HP 1 mL nickel affinity column, and washed with washing buffer (10 mL, section 7.2.5). Protein was eluted with buffers of increasing imidazole concentration (3 mL each - 20 mM Tris-HCl, 100 mM NaCl, 10 % glycerol, and 100/200/300 mM imidazole).

An 10% acrylamide/bis-acrylamide SDS-PAGE gel was prepared (table 7.3), adding the stacking gel with comb to the loading gel once it had set. When the entire gel was set the comb was removed and the gel placed in a BioRad Mini PROTEAN Tetra Cell tank, submerging in running buffer (2 M glycine, 1% SDS, 250 mM Tris-HCl, pH 8). Protein fractions (15 μL) were mixed with SDS-PAGE loading dye (5 μL, 1.5 mL 1 M Tris-HCl pH 6.8, 3 mL 1 M dithiothreitol – DTT, 0.6 g SDS, 0.03 g bromophenol blue, 2.4 mL glycerol and deionized water to a total volume of 7.5 mL), and loaded onto to the SDS-PAGE gel along with Thermo Scientific Page Ruler Plus Prestained Protein Ladder (6 μL). Electrophoresis was run at 180 V for ca. 45 minutes using a BioRad Power Pac 300 power supply, and the protein bands visualized using an Expedeon Instant Blue stain.

Table 7.3: Reagents and quantities required for a 10% acrylamide/bis-acrylamide SDS-PAGE gel. APS = Ammonium persulfate.

Reagents	10% Loading gel (μL)	Stacking gel (μL)
Deionized water	2000	1050
30% acrylamide/bis-acrylamide	1650	250
1.5 m Tris-HCl pH 8.8	1250	-
1.0 m Tris-HCl pH 6.8	-	190

10% SDS	50	15
10% APS	50	15
TEMED	2.0	1.5

Bands corresponding to an appropriate mass for the proteins were excised and trypsin digested using a Sigma Aldrich Trypsin Profile IGD kit as per the manufacturer's specifications, and the peptide mass fingerprint analysed using MALDI-TOF mass spectrometry using a Sigma Aldrich α-cyano-4-hydroxycinnamic acid matrix and a Bruker Autoflex Spectrometer.

TmcI was loaded onto a GE Healthcare PD-10 column, eluting with concentration buffer (20 mM Tris-HCl, pH 8, 100 mM NaCl, 10% glycerol), in order to remove excess imidazole. Proteins were concentrated by centrifugation using a GE Healthcare 30 kDa MWCO Vivaspin 20 spin filter, washing with concentration buffer (2 × 5 mL). Concentrations of the proteins was determined using a Thermo Scientific NanoDrop Lite spectrophotometer, and the proteins divided into aliquots (20 µL each) and stored at -80 °C. ESI-MS analyses were conducted using a Bruker MaXis Impact spectrometer connected to a Dionex 3000 RS UHPLC instrument fitted with a Chromatography Technologies ACE 3 C4-300 reverse phase column (100 × 2.1 mm, 1.8 µm, 25 °C).

TmcI yield = 1.8 mg/L

EpnF yield = 11.1 mg/L

7.2.7 Characterization of heme cofactor for TmcI

TmcI was analysed by UHPLC-ESI-Q-TOF MS using the following elution profile: 0 min: 5% acetonitrile + 0.1 % formic acid, 95% water + 0.1 % formic acid; 30 min: 5% acetonitrile + 0.1 % formic acid ; 40 min: 5% acetonitrile + 0.1 % formic acid, using a Bruker MaXis Impact spectrometer connected to a Dionex 3000 RS UHPLC instrument fitted with a Chromatography Technologies ACE 3 C4-300 reverse phase column (100 × 2.1 mm, 1.8 µm, 25 °C). An absorbance for the heme prosthetic group

was observed in the 280 nm UV chromatogram, as well as in the corresponding EIC (*m/z* calcd for $C_{34}H_{32}FeN_4O_4^+$: 616.1767, observed: 616.1765).

The purity index heme content of desalted TmcI (8 M) in Tris-HCl, pH 8 (25 mM) was evaluated by measuring absorbance between 500-250 nm using a Varian Cary 50 Bio UV-Visible spectrophotometer. A purity index of 1.0 was calculated from the ratio A_{418}/A_{280}.[72]

7.2.8 Characterization of a flavin cofactor for EpnF

The UV-Vis absorbance of EpnF (20 μM) in concentration buffer (section 7.2.6) was measured between 500-300 nm using a Varian Cary 50 Bio UV-Visible spectrophotometer. An absorbance maximum at 447 nm was observed, consistent with a bound flavin cofactor.

7.2.9 Oligomerization state of EpnF and TmcI

A calibration solution of thyroglobulin (bovine thyroid, Sigma), apoferritin (equine spleen, Sigma), β-amylase (sweet potato, Sigma), bovine serum albumin (bovine, Sigma) and lysozyme (hen egg white, Fluka) in concentration buffer (200 μL, section 7.2.6) was analysed by gel filtration chromatography on a GE Healthcare Superdex 200 column using concentration buffer as the eluent and monitoring absorbance at 280 nm (table 7.4). Comparison of the elution volumes for TmcI and EpnF with those of the molecular weight lead to the conclusion that they are monomeric and dimeric proteins respectively.[42]

Table 7.4: Molecular weights and elution volumes for proteins used to calibrate the gel filtration column.

Protein	MWt (kDa)	Elution volume (mL)
Thyroglobulin	669	46.6
Apoferritin	443	50.4
β-Amylase	200	65.5

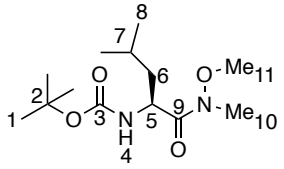
Bovine serum albumin	66	76.1
Lysozyme	13	108.6

7.3. Chemical synthesis

All compounds synthesised have been characterised by ^{1}H and ^{13}C NMR analyses. Novel compounds were additionally characterised by HR-MS analysis, while known compounds were characterised by MS analysis.

7.3.1 Synthesis of isopropyl ketone putative substrate

96 *tert*-Butyl *(S)*-(1-(methoxy(methyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate



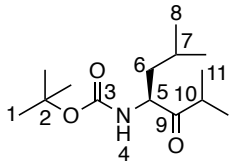
To a stirred solution of BOC-L-leucine (2.5 g, 11.0 mmol) in DCM (50 mL) was added *O*,*N*-dimethylhydroxylamine hydrochloride (1.2 g, 12.1 mmol), triethylamine (3.2 mL, 23.0 mmol) and BOP reagent (4.9 g, 11.0 mmol). After 3.5 hours the solution was diluted with DCM (200 mL) and washed with HCl (3 M, 3 × 30 mL), saturated $NaHCO_3$ (3 × 30 mL) and brine (3 × 30 mL). The organic layer was dried ($MgSO_4$), solvent removed in *vacuo* and the resulting residue purified by silica gel chromatography (1:3 ethyl acetate:*n*-hexane) to give the desired product as a yellow oil (2.3 g, 8.5 mmol, 77%).[22]

δ_H ($CDCl_3$, 400 MHz): 5.04 (1H, br d, *J* 9.0 Hz, H-4), 4.77-4.62 (1H, m, H-5), 3.77 (3H, s, H-11), 3.18 (3H, s, H-10), 1.70 (1H, septet, *J* 7.0 Hz, H-7), 1.48-1.36 (11H, m, H-1, H-6), 0.95 (3H, d, *J* 6.5 Hz, H-8), 0.91 (3H, d, *J* 6.5 Hz, H-8).

δ_C ($CDCl_3$, 100 MHz): 177.2 (C-9), 155.8 (C-3), 79.6 (C-2), 61.7 (C-11), 49.1 (C-5), 42.2 (C-6), 32.2 (C-10), 28.5 (C-1), 24.8 (C-7), 23.5 (C-8), 21.7 (C-8).

m/z ES^+ (%): 275 (9, $[M+H]^+$), 297 (100, $[M+Na]^+$).

100 *tert*-Butyl *(S)*-(2,6-dimethyl-3-oxoheptan-4-yl)carbamate



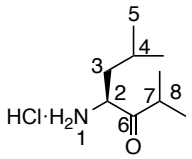
Under an argon atmosphere, a stirred solution of **96** (2.3 g, 8.5 mmol) in THF (13 mL) was cooled to 0 °C and isopropylmagnesium chloride (2 M in diethyl ether, 21.5 mL, 43.0 mmol) was added dropwise. The reaction was allowed to reach 25 °C and, after 4 hours stirring, was quenched with HCl (1 M, 2.5 mL), forming a precipitate. The precipitate was removed by filtration and washed with ethyl acetate (3 × 50 mL). The organic extractions were combined and washed with brine (3 × 30 mL). The organic layer was dried ($MgSO_4$), solvent removed in *vacuo* and the resulting residue purified by silica gel chromatography (1:4 ethyl acetate:*n*-hexane) to give the desired product as a colorless oil (1.1 g, 4.3 mmol, 51%).[101]

δ_H ($CDCl_3$, 400 MHz): 5.00 (1H, br d, *J* 8.0 Hz, H-4), 4.49 (1H, dt, *J* 9.5, 3.5 Hz, H-5), 2.80 (1H, septet, *J* 7.0 Hz, H-10), 1.77-1.65 (1H, m, H-7), 1.53-1.39 (10H, m, H-1, H-6), 1.34-1.25 (1H, m, H-6), 1.14 (3H, d, *J* 7.0 Hz, H-11), 1.11 (3H, d, *J* 7.0 Hz, H-11), 0.98 (3H, d, *J* 6.5 Hz, H-8), 0.93 (3H, d, *J* 6.5 Hz, H-8).

δ_C ($CDCl_3$, 100 MHz): 213.9 (C-9), 155.5 (C-3), 79.3 (C-2), 56.0 (C-5), 40.7 (C-6), 37.6 (C-10), 28.2 (C-1), 24.8 (C-7), 23.3 (C-8), 21.6 (C-8), 18.8 (C-11), 17.7 (C-11).

m/z ES^+ (%): 256 (5, $[M+H]^+$), 280 (100, $[M+Na]^+$).

94 *(S)*-2,6-Dimethyl-3-oxoheptan-4-ammonium chloride[42]



To **100** (1.1 g, 4.3 mmol) at 0 °C was added HCl in 1,4-dioxane (4 M, 86 mL). The resulting mixture was stirred while warming to 25 °C. After a further 45 minutes

stirring, the solvent was removed in *vacuo*, and the resulting residue washed with diethyl ether (3 × 10 mL) to yield the desired product as a brown oil (459 mg, 2.4 mmol, 55%).

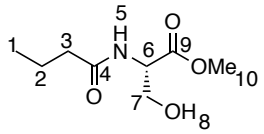
δ_H ($CDCl_3$, 400 MHz): 8.52 (3H, br s, H-1), 4.42 (1H, s, H-2), 2.82 (1H, septet, *J* 6.5 Hz, H-7), 2.13-2.01 (1H, m, H-4), 1.93-1.83 (1H, m, H-3), 1.67-1.57 (1H, m, H-3), 1.20 (3H, d, *J* 7.5 Hz, H-8), 1.11 (3H, d, *J* 7.5 Hz, H-8), 1.04 (3H, d, *J* 6.0 Hz, H-5), 0.99 (3H, d, *J* 6.5 Hz, H-5).

δ_C ($CDCl_3$, 100 MHz): 209.7 (C-6), 56.4 (C-2), 39.1 (C-3), 37.5 (C-7), 24.8 (C-4), 23.4 (C-5), 21.6 (C-5), 19.3 (C-8), 17.6 (C-8).

HRMS (ESI^+): calcd for $C_9H_{20}NO^+$: 158.1539, found: 158.1535.

IR (cm^{-1}): 2936 (N-H stretch), 1719 (C=O stretch).

104 Methyl butyryl-L-serinate[42]



Under an argon atmosphere to a stirred solution of L-serine methyl ester hydrochloride (2.0 g, 13.0 mmol) in DCM (10 mL) was added triethylamine (1.8 mL, 12.9 mmol) at -20 °C. After stirring for 10 minutes to this was added freshly distilled butyric acid (1.0 mL, 10.9 mmol) and DCC (2.2 g, 10.9 mmol). The solution was allowed to reach 25 °C and left to stir for 16 hours. Urea precipitate was removed by filtration and solvent removed in *vacuo*. The resulting residue was purified by silica gel chromatography (ethyl acetate) to give the desired product as a white powder (1.3 g, 7.1 mmol, 65%).

δ_H ($CDCl_3$, 400 MHz): 6.51 (1H, br d, *J* 6.0 Hz, H-5), 4.67 (1H, quintet, *J* 4.0 Hz, H-6), 3.93 (1H, dd, *J* 11.0, 3.5 Hz, H-7), 3.89 (1H, dd, *J* 11.0, 3.5 Hz, H-7), 3.78 (3H, s, H-10), 2.99 (1H, br s, H-8), 2.24 (2H, t, *J* 7.0, H-3), 1.68 (2H, sextet, *J* 7.5 Hz, H-2), 0.96 (3H, t, *J* 7.5 Hz, H-1).

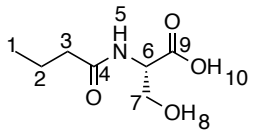
δ_C ($CDCl_3$, 100 MHz): 173.8 (C-4), 171.2 (C-9), 63.6 (C-7), 54.7 (C-6), 52.9 (C-10), 38.5 (C-3), 19.1 (C-2), 13.8 (C-1).

HRMS (ESI$^+$): calcd for $C_8H_{15}NNaO_4^+$: 212.0893, found: 212.0891.

IR (cm^{-1}): 3294 (O-H stretch), 1740 (C=O stretch).

$[\alpha]_D^{20}$ = +13.5 (*c* 0.375, $CHCl_3$).

93 *N*-Butanoyl-L-serine[42]



To a stirred solution of **104** (1.3 g, 7.1 mmol) in THF (17 mL) was added 30% (*w/w*) hydrogen peroxide (0.17 mL) and aqueous lithium hydroxide (0.76 M, 14.2 mL, 10.7 mmol) dropwise. After 2 hours the solvent was removed in *vacuo* and the residue dissolved in deionized water (10 mL). The aqueous layer was washed with diethyl ether (2 × 10 mL), acidified to pH 1 with hydrochloric acid (1 M), and sodium chloride added (ca. 5 g). The aqueous layer was extracted with ethyl acetate (20 × 10 mL), and the organics combined and dried ($MgSO_4$). Solvent was removed in *vacuo* to give the desired product as a clear oil (969 mg, 5.5 mmol, 78%).

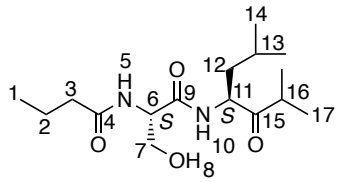
δ_H (MeOD, 300 MHz): 4.50 (1H, br t, *J* 4.5 Hz, H-6), 3.90 (1H, dd, *J* 11.0, 5.0 Hz, H-7), 3.82 (1H, dd, *J* 11.0, 4.0 Hz, H-7), 2.26 (2H, t, *J* 7.5 Hz, H-3), 1.66 (2H, sextet, *J* 7.5 Hz, H-2), 0.97 (3H, t, 7.5 Hz, H-1).

δ_C ($CDCl_3$, 75 MHz): 176.2 (C-4), 173.5 (C-9), 63.0 (C-7), 56.1 (C-6), 38.7 (C-3), 20.2 (C-2), 14.0 (C-1).

HRMS (ESI$^-$): calcd for $C_7H_{12}NO_4^-$: 174.0772, found: 174.0772.

IR (cm^{-1}): 3308 (O-H stretch), 1720 (C=O stretch).

71 *N-((S)*-1-*(((S)*-2,6-Dimethyl-3-oxoheptan-4-yl)amino)-3-hydroxy-1-oxopropan-2-yl)butyramide[42]



To a stirred solution of **104** (31.7 mg, 181 μmol) in dioxane (1.3 mL) was added triethylamine (25 μL, 179 μmol), BOP reagent (81.1 mg, 183 μmol) and a solution of **94** (38.5 mg, 199 μmol) in dioxane (1.7 mL) and triethylamine (28 μL, 200 μmol). After 4 hours, precipitate formed was removed by filtration and solvent removed in *vacuo*. Resultant residue was dissolved in ethyl acetate (10 mL), then washed with HCl (3 M, 3 × 10 mL), saturated $NaHCO_3$ (3 × 10 mL) and brine (3 × 10 mL). The organic layer was dried ($MgSO_4$), solvent removed in *vacuo* and resulting residue purified by silica gel chromatography (ethyl acetate) to yield a mixture of diastereomers (demonstrated to have distinct retention times by LCMS) as a light brown solid (14.9 mg, 47 μmol, 26%, major:minor 8:5).

HRMS (ESI^+): calcd for $C_{16}H_{30}N_2NaO_4^+$: 337.2098: found 337.2100.

IR (cm^{-1}): 3313 (O-H stretch), 1715 (C=O stretch).

Predicted major diastereomer (*SS*):

δ_H ($CDCl_3$, 500 MHz): 7.24-7.19 (1H, m, H-10), 6.68 (1H, br d, *J* 7.0 Hz, H-5), 4.74-4.69 (1H, m, H-11), 4.57-4.41 (1H, m, H-6), 4.10-4.04 (1H, m, H-7), 3.70-3.53 (1H, m, H-7), 2.86-2.74 (1H, septet, *J* 6.5 Hz, H-16), 2.25 (2H, t, *J* 7.0 Hz, H-3), 1.72-1.62 (3H, m, H-2, H-13), 1.57-1.50 (1H, m, H-12), 1.45-1.37 (1H, m, H-12), 1.13 (3H, d, *J* 7.0 Hz, H-17), 1.10-1.06 (3H, m, H-17), 0.98-0.91 (9H, m, H-1, H-14).

δ_C ($CDCl_3$, 125 MHz): 212.6 (C-15), 173.3 (C-4), 170.1 (C-9), 62.1 (C-7), 54.5 (C-11), 53.1 (C-6), 39.1 (C-3), 37.4 (C-12), 36.8 (C-16), 24.3 (C-13), 22.5 (C-14), 20.6 (C-14), 18.2 (C-17), 18.1 (C-2), 16.9 (C-17), 12.8 (C-1).

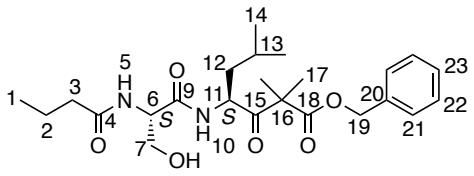
Predicted minor diastereomer (*SR*):

δ_H ($CDCl_3$, 500 MHz): 7.15-7.10 (1H, m, H-10), 6.59 (1H, br d, *J* 7.0 Hz, H-5), 4.57-4.51 (1H, m, H-6), 4.74-4.69 (1H, m, H-11), 4.01 (1H, dd, *J* 12.0, 3.5 Hz, H-7), 3.88-3.79 (1H, m, H-7), 2.86-2.74 (1H, septet, *J* 6.5 Hz, H-16), 2.21 (2H, dt, *J* 7.5, 2.0 Hz, H-3), 1.72-1.62 (3H, m, H-2, H-13), 1.57-1.50 (1H, m, H-12), 1.45-1.37 (1H, m, H-12), 1.13 (3H, d, *J* 7.0 Hz, H-17), 1.10-0.96 (3H, m, H-17), 0.98-0.91 (9H, m, H-1, H-14).

δ_C ($CDCl_3$, 125 MHz): 212.1 (C-15), 172.8 (C-4), 170.2 (C-9), 62.1 (C-7), 55.0 (C-11), 52.9 (C-6), 38.8 (C-3), 37.4 (C-12), 36.7 (C-16), 24.3 (C-13), 22.5 (C-14), 20.5 (C-14), 18.3 (C-17), 18.1 (C-2), 16.9 (C-17), 12.8 (C-1).

7.3.2 Synthesis of α-dimethyl-β-keto carboxylic acid putative substrate

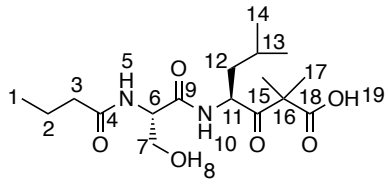
123 Benzyl (*S*)-4-((*S*)-2-butyramido-3-hydroxypropanamido)-2,2,6-trimethyl-3-oxoheptanoate[42]



To a solution of **93** (42.1 mg, 240 μmol) in dichloromethane (DCM, 3 mL) and triethylamine (36.5 μL, 262 μmol) was added (*S*)-1-(Benzyloxy)-2,2,6-trimethyl-1,3-dioxoheptan-4-ammonium trifluoroacetate (96.6 mg, 238 μmol), HOBT (68.2 mg, 505 μmol) and EDC (91.4 mg, 477 μmol). After 24 hours stirring at 25 °C the solution was washed with saturated $NaHCO_3$ (3 × 5 mL), deionized water (5 mL) and brine (5 mL), and the organic layer was dried ($MgSO_4$). The solvent was removed in *vacuo* and the resulting residue purified by silica gel chromatography (1:1 ethyl acetate:*n*-hexane) to yield the desired product as a yellow crystalline solid (27.6 mg, 62 μmol, 26%). ^{1}H NMR spectroscopic analysis indicated that it contained approximately 10% of the C-11 epimer. The data reported are for predicted major diastereomer (*SS*).

δ_H ($CDCl_3$, 400 MHz): 7.59-7.24 (5H, m, H-21, H-22, H-23), 7.06 (1H, d, *J* 8.5 Hz, H-10), 6.61 (1H, d, *J* 6.5 Hz, H-5), 5.24-5.09 (2H, m, H-19), 4.96-4.88 (1H, m, H-11), 4.40-4.34 (1H, m, H-6), 3.97 (1H, dd, *J* 11.5, 2.5 Hz, H-7), 3.53 (1H, dd, *J* 11.5, 5.0 Hz, H-7), 2.21 (2H, dt, *J* 7.5, 3.0 Hz, H-3), 1.71-1.60 (2H, m, H-2), 1.58-1.29 (3H, m, H-12, H-13), 1.45 (3H, s, H-17), 1.41 (3H, s, H-17), 0.97-0.82 (9H, m, H-1, H-14).
δ_C ($CDCl_3$, 100 MHz): 207.5 (C-15), 174.1 (C-4), 173.1 (C-18), 171.0 (C-9), 135.4 (C-20), 128.8 (C-21), 128.6 (C-23) 128.3 (C-22), 67.5 (C-19), 62.7 (C-7), 55.0 (C-16), 53.6 (C-6) 53.3 (C-11), 41.1 (C-12), 38.3 (C-3), 24.9 (C-13), 23.5 (C-14), 22.6 (C-17), 22.4 (C-17), 21.2 (C-14), 19.1 (C-2), 13.8 (C-1).
HRMS (ESI^+): calculated for $C_{24}H_{36}N_2NaO_6^+$: 471.2466, found: 471.2465.

66 *(S)*-4-*((S)*-2-Butyramido-3-hydroxypropanamido)-2,2,6-trimethyl-3-oxoheptanoic acid[42]



123 (27.6 mg, 62 μmol) was stirred in methanol (2 mL) containing 10% Pd/C (ca. 25 mg) under a hydrogen atmosphere (1 atm). After 2 hours the mixture was filtered through celite and the solvent was removed in *vacuo* to give the product as a clear oil (16.3 mg, 47 μmol, 75%). ^{1}H NMR spectroscopic analysis showed that this contained approximately 20% of **71** resulting from decarboxylation of the product. Further decarboxylation was observed to occur during subsequent spectroscopic analysis. After 48 hours the ratio of α-dimethyl-β-keto acid to isopropyl ketone was estimated to be 4:5 by ^{1}H NMR spectroscopic analysis. Spectroscopic data for the *α*-dimethyl-*β*-keto acid are given below. Spectroscopic data for **71** are reported in the experimental section for that compound.[42]
δ_H ($CDCl_3$, 500 MHz): 7.15-7.10 (1H, br s, H-10), 6.67 (1H, br s, H-5), 5.01-4.94 (1H, m, H-11), 4.75-4.69 (1H, m, H-6), 4.03-3.92 (1H, m, H-7), 3.64-3.59 (1H, m, H-7),

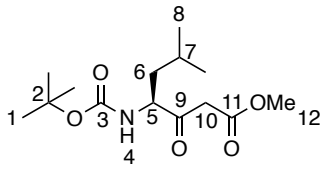
2.26 (2H, t, *J* 6.5 Hz, H-3), 1.72-1.61 (1H, m, H-13), 1.60-1.50 (2H, m, H-12), 1.41 (3H, s, H-17), 1.39 (3H, s, H-17), 0.98- 0.89 (H11, m, H-14, H-1, H-2).

δ_C ($CDCl_3$, 125 MHz): 207.6 (C-15), 175.9 (C-18), 170.8 (C-9), 168.1 (C-4), 63.9 (C-7), 54.8 (C-16), 53.9 (C-11), 53.5 (C-6), 38.3 (C-3), 24.8 (C-13), 22.5 (C-14), 21.8 (C-14), 19.2 (C-2), 13.8 (C-1).

HRMS (ESI^-): calculated for $C_{17}H_{29}N_2O_6^-$: 357.2031, found: 357.2028.

7.3.3 Synthesis of *α*-dimethyl-*β*-keto methyl ester substrate analogues

143 Methyl *(S)*-4-((BOC)amino)-6-methyl-3-oxoheptanoate



Method A: To a solution of *N*-BOC-L-leucine (1.5 g, 6.5 mmol) in DCM (20 mL) was added EDC (1.4 g, 7.1 mmol), DMAP (873 mg, 7.1 mmol) and Meldrum's acid (recrystallized from acetone prior to use, 937 mg, 6.5 mmol). Reaction was stirred and monitored by MS, revealing the reaction had not gone to completion after 27 hours, additional EDC (496 mg, 2.6 mmol) and DMAP (320 mg, 2.6 mmol) were added, and the reaction stirred for an additional 69 hours. The reaction was mixture was poured in HCl (1 M, 20 mL) and the organic layer collected. The aqueous layer was extracted with DCM (3 × 20 mL), the organic extractions combined and dried ($MgSO_4$). Solvent was removed in *vacuo* to provide the Meldrum's adduct as a yellow solid (1.8 g, 5.1 mmol, ca. 79%), which was used without further purification.

Meldrum's adduct (1.8 g, 5.1 mmol) was dissolved in methanol (20 mL) and refluxed at 70 °C for 24 hours. Solvent was removed in *vacuo,* and the resulting residue was purified by silica gel chromatography (2:3 ethyl acetate:*n*-hexane) to give the desired product as a yellow oil (573 mg, 2.0 mmol, 31%).

Method B: Under an atmosphere of argon to a stirred solution *N*-BOC-L-leucine (1.0 g, 4.3 mmol) in THF (9.9 mL) was added CDI (841 mg, 5.2 mmol). The solution was stirred at room temperature for two hours. Methyl potassium malonate (1.0 g, 6.5 mmol) and $MgCl_2$ (624 mg, 5.2 mmol) were added and the reaction allowed to stir overnight. Solvent was removed in *vacuo*, and the crude residue was partitioned between ethyl acetate (40 mL) and HCl (1 M, 40 mL). The organic layer was collected, and the aqueous phase was extracted with ethyl acetate (3 × 40 mL). Organic extractions were combined, washed with 5% $NaHCO_3$ (*w/w*, 2 × 40 mL) and brine (40 mL), and dried ($MgSO_4$). Solvent was removed in *vacuo* to provide the desired product as a yellow oil (718 mg, 2.5 mmol, 58%).[86]

δ_H ($CDCl_3$, 300 MHz): 4.91 (1H, br d, *J* 8.0 Hz, H-4), 4.40-4.26 (1H, m, H-5), 3.73 (3H, s, H-12), 3.56 (2H, d, *J* 6.5 Hz, H-10), 1.78-1.54 (2H, m, H-6, H-7), 1.43 (9H, s, H-1), 1.41-1.33 (1H, m, H-6), 0.97-0.91 (6H, m, H-8).

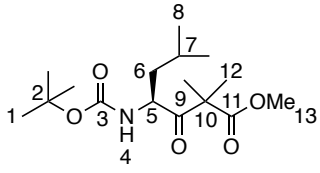
δ_C ($CDCl_3$, 75 MHz): 202.9 (C-9), 167.6 (C-11), 162.5 (C-3), 80.3 (C-2), 58.3 (C-5), 52.5 (C-12), 46.1 (C-10), 39.9 (C-6), 28.4 (C-1), 24.9 (C-7), 23.3 (C-8), 21.7 (C-8).

m/z ES^+ (%): 310 (100, $[M+Na]^+$).

IR (cm^{-1}): 1751 (C=O stretch), 1704 (C=O stretch).

$[\alpha]_D^{20}$ = -16.1 (*c* 0.375, $CHCl_3$).

144 Methyl *(S)*-4-((BOC)amino)-2,2,6-methyl-3-oxoheptanoate



To a stirred solution of **143** (189 mg, 658 μmol) and potassium carbonate (909 mg, 6.6 mmol) under reflux (70 °C) and an atmosphere of argon was added iodomethane (400 μL, 6.4 mmol). The reaction was allowed to stir for 23 hours. Solvent was removed in *vacuo* and the crude residue partitioned between water (20 mL) and ethyl acetate (20 mL). Sodium chloride (ca. 1 g) was added to the mixture, and the organic layer collected. The aqueous layer was extracted with

ethyl acetate (3 × 20 mL), the organic layers combined and washed with brine (2 × 40 mL), and dried ($MgSO_4$). Solvent was removed in *vacuo*, and the resulting residue was purified with was purified by silica gel chromatography (1:4 ethyl acetate:*n*-hexane) to give the desired product as a clear oil (82.4 mg, 261 μmol 40%).[102]

δ_H ($CDCl_3$, 300 MHz): 4.78 (1H, br d, *J* 10.0 Hz, H-4), 4.65, (1H, dt, *J* 10.0, 4.0 Hz, H-5), 3.73 (3H, s, H-13), 1.75-1.62 (H-7), 1.45-1.40 (15H, m, H-1, H-12), 1.39-1.30 (2H, m, H-6), 0.95 (3H, d, *J* 6.5 Hz), 0.92 (3H, d, *J* 6.5 Hz).

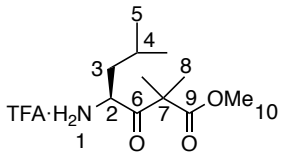
δ_C ($CDCl_3$, 75 MHz): 209.0 (C-9), 173.0 (C-11), 162.5 (C-3), 79.8 (C-2), 54.3 (C-10), 54.0 (C-5), 52.7 (C-13), 41.9 (C-6), 28.4 (C-1), 24.8 (C-7), 23.7 (C-8), 22.5 (C-12), 22.2 (C-12), 21.5 (C-8).

m/z ES^+ (%): 338 (100, $[M+Na]^+$).

IR (cm^{-1}): 1743 (C=O stretch), 1706 (C=O stretch).

$[\alpha]_D^{20}$ = -42.9 (*c* 0.275, $CHCl_3$).

145 *(S)*-1-Methoxy-2,2,6-trimethyl-1,3-dioxoheptan-4-ammonium trifluoroacetate



Under an atmosphere of argon to a 10% TFA in DCM solution (5 mL) while stirring was added **144** (74.7 mg, 236 μmol). After half an hour the solution was diluted with DCM (3 × 15 mL), removing solvent in *vacuo* successively. The resulting crude oil was washed with petroleum ether (3 × 5 mL), and residual solvent removed in *vacuo* from the immiscible oil to give the desired product as a brown oil (74.6 mg, 227 μmol, 96%).

δ_H (MeOD, 500 MHz): 4.52-4.48 (1H, m, H-2), 3.76 (3H, s, H-10), 1.84-1.69 (1H, m, H-4), 1.61-1.55 (2H, m, H-3), 1.48 (3H, s, H-8), 1.47 (H-8), 1.04-0.98 (6H, m, H-5).

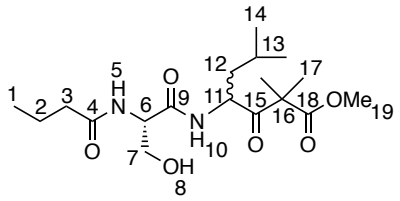
δ_C (MeOD, 125 MHz): 206.1 (C-6), 174.5 (C-9), 55.7 (C-7), 55.6 (C-2), 53.4 (C-10), 40.5 (C-3), 25.6 (C-4), 23.7 (C-5), 23.1 (C-8), 22.9 (C-8), 20.8 (C-5).

HRMS (ESI$^+$): calcd for $C_{11}H_{22}NO_3^+$: 216.1594, found: 216.1597.

IR (cm^{-1}): 2935 (N-H stretch), 1763 (C=O stretch).

$[\alpha]_D^{20}$ = -5.7 (*c* 0.290, $CHCl_3$).

139 Methyl 4-*((S)*-2-butyramido-3-hydroxypropanamido)-2,2,6-trimethyl-3-oxoheptanoate



To a stirred solution of **93** (39.7 mg, 219 µmol) in DMF (1 mL) was added HOBt (62.0 mg, 459 µmol), EDC (84.4 mg, 440 µmol) and triethylamine (23.0 µL, 165 µmol), and the solution was stirred for 1 hour. **145** (72.4 mg, 231 µmol) in DMF (1 mL) with triethylamine (23.0 µL, 165 µmol) was added, and the reaction stirred for a further 90 hours. Reaction mixture was diluted with ethyl acetate (5 mL), and washed with HCl (1 M, 5 mL), $NaHCO_3$ (1 M, 5 mL) and brine (3 × 5 mL). Organic layer was dried ($MgSO_4$) and solvent removed in *vacuo* to provide crude residue (36.3 mg). To obtain more material, aqueous washings were further extracted with ethyl acetate (3 × 10 mL), organic extractions combined and washed with brine (10 mL), and dried ($MgSO_4$). Organic extractions were combined with original crude residue and solvent was removed in *vacuo*. The resulting residue was purified with was purified by silica gel chromatography (a gradient of 1% methanol in DCM to 10% methanol in DCM) to yield the product as a brown oil (17.7 mg, 47.5 µmol, 22%). ^{1}H NMR spectroscopic analysis indicated the compound was approximately a 1:1 mixture of diastereomers, which were assigned without evaluating their stereochemistry.

HRMS (ESI$^+$): calcd for $C_{18}H_{32}N_2NaO_6^+$: 395.2153, found: 395.2152.

IR (cm^{-1}): 3271 (O-H stretch), 1761 (C=O stretch), 1744 (C=O stretch), 1693 (C=O stretch).

Diastereomer 1:

δ_H (MeOD, 500 MHz): 4.99-4.91 (1H, m, H-11), 4.44-4.38 (1H, m, H-6), 3.79-3.67 (2H, m, H-7), 3.71 (3H, s, H-19), 2.24 (2H, br t, *J* 7.5 Hz, H-3), 1.69-1.60 (3H, m, H-2, H-13), 1.56-1.41 (2H, m, H-12), 1.40 (3H, s, H-17), 1.36 (3H, s, H-17), 0.98-0.89 (9H, m, H-1, H-14).

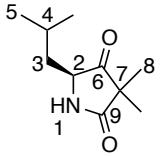
δ_C (MeOD, 125 MHz): 208.5 (C-11), 176.3 (C-4), 175.0 (C-18), 172.0 (C-9), 62.9 (C-7), 56.6 (C-6), 56.2 (C-16), 53.6 (C-19), 53.2 (C-11), 42.0 (C-12), 38.7 (C-3), 25.7 (C-13), 23.8 (C-14), 22.6 (C-17), 21.7 (C-14), 20.2 (C-2), 14.0 (C-1).

Diastereomer 2:

δ_H (MeOD, 500 MHz): 4.99-4.91 (1H, m, H-11), 4.44-4.38 (1H, m, H-6), 3.79-3.67 (2H, m, H-7), 3.70 (3H, s, H-19), 2.24 (2H, br t, *J* 7.5 Hz, H-3), 1.69-1.60 (3H, m, H-2, H-13), 1.56-1.41 (2H, m, H-12), 1.40 (3H, s, H-17), 1.36 (3H, s, H-17), 0.98-0.89 (9H, m, H-1, H-14).

δ_C (MeOD, 125 MHz): 208.3 (C-11), 176.1 (C-4), 174.9 (C-18), 171.9 (C-9), 63.0 (C-7), 56.5 (C-6), 56.2 (C-16), 53.6 (C-19), 53.1 (C-11), 42.0 (C-12), 38.7 (C-3), 25.7 (C-13), 23.8 (C-14), 22.6 (C-17), 21.7 (C-14), 20.2 (C-2), 14.0 (C-1).

151 *(S)*-5-isobutyl-3,3-dimethylpyrrolidine-2,4-dione



Under an atmosphere of argon to a stirred solution of **93** (79.5 mg, 439 μmol), **145** (139 mg, 443 μmol) and HOBt (143 mg, 1.1 mmol) in THF (8.9 mL) at 0 °C was added triethylamine (260 μL, 1.9 mmol) and EDC (173 mg, 902 μmol). The reaction was warmed to room temperature and stirred for 24 hours. Reaction mixture was filtered, and solvent removed in *vacuo*. The resulting residue was purified by silica

gel chromatography (ethyl acetate) to provide the lactam byproduct **151** as a white solid (29.4 mg, 161 μmol, 36%) and none of the desired **139**.

δ_H (MeOD, 500 MHz): 4.09 (1H, dd, *J* 9.0, 5.0 Hz, H-2), 1.88-1.76 (1H, m, H-4), 1.63-1.55 (1H, m, H-3), 1.50-1.42 (1H, m, H-3), 1.20 (3H, s, H-8), 1.17 (3H, s, H-8), 0.97 (3H, d, *J* 4.0 Hz, H-5), 0.95 (3H, d, *J* 4.0 Hz, H-5).

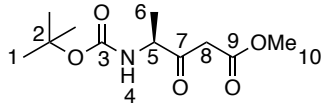
δ_C (MeOD, 125 MHz): 216.3 (C-6), 180.3 (C-9), 61.1 (C-2), 43.1 (C-3), 25.9 (C-4), 23.6 (C-5), 21.8 (C-5), 21.3 (C-8), 21.2 (C-8). The signal due to C-7 was not observed due to low relative intensity.

HRMS (ESI^+): calcd for $C_{10}H_{17}NNaO_2^+$: 206.1151, found: 206.1152.

IR (cm^{-1}): 1765 (C=O stretch), 1692 (C=O stretch).

$[\alpha]_D^{20}$ = -73.0 (*c* 0.265, $CHCl_3$).

156 Methyl *(S)*-4-((BOC)amino)-3-oxopentanoate



Under an atmosphere of argon to a stirred solution *N*-BOC-L-alanine (2.0 g, 10.6 mmol) in THF (24.4 mL) was added CDI (2.1 g, 12.7 mmol). The solution was stirred at room temperature for two hours. Methyl potassium malonate (2.5 g, 16.0 mmol) and $MgCl_2$ (1.5 g, 12.7 mmol) were added and the reaction allowed to stir overnight. Solvent was removed in *vacuo*, and the crude residue was partitioned between ethyl acetate (40 mL) and HCl (1 M, 40 mL). The organic layer was collected, and the aqueous phase was extracted with ethyl acetate (3 × 40 mL). Organic extractions were combined, washed with 5% $NaHCO_3$ (*w/w*, 2 × 40 mL) and brine (40 mL), and dried ($MgSO_4$). Solvent was removed in *vacuo* to yield the desired product as a pale yellow solid (2.2 g, 8.9 mmol, 84%).[103]

δ_H ($CDCl_3$, 300 MHz): 5.09 (1H, br s), 4.38 (1H, br t, *J* 7.5 Hz), 3.74 (3H, s), 3.57 (1H, d, *J* 16.0 Hz), 3.54 (1H, d, *J* 16.0 Hz), 1.45 (9H, s), 1.36 (3H, d, *J* 7.0 Hz).

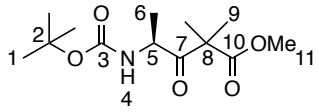
δ_C ($CDCl_3$, 75 MHz): 202.3 (C-7), 167.7 (C-9), 162.5 (C-3), 80.1 (C-2), 55.5 (C-5), 52.7 (C-10), 45.7 (C-8), 28.5 (C-1), 17.1 (C-6).

m/z ES^+ (%) 268 (100, $[M+Na]^+$).

IR (cm^{-1}): 1749 (C=O stretch), 1718 (C=O stretch), 1683 (C=O stretch).

$[\alpha]_D^{20}$ = -18.6 (*c* 0.350, $CHCl_3$).

157 Methyl *(S)*-4-((BOC)amino)-2,2-dimethyl-3-oxopentanoate



Under an atmosphere of argon to a stirred solution **157** (300 mg, 1.2 mmol) and potassium carbonate (1.7 g, 12.3 mmol) in THF (10.2 mL) was added iodomethane (760 µL, 12.2 mmol), and the solution stirred for 17.5 hours. Solvent was removed in *vacuo* and the crude residue partitioned between water (20 mL) and ethyl acetate (20 mL). Sodium chloride (ca. 1 g) was added to the mixture, and the organic layer collected. The aqueous layer was extracted with ethyl acetate (3 × 20 mL), the organic extractions combined and washed with brine (2 × 40 mL), and dried ($MgSO_4$). Solvent was removed in *vacuo*, and the resulting residue was purified with was purified by silica gel chromatography (3:17 ethyl acetate:*n*-hexane) to give the desired product as pale yellow solid (244 mg, 893 µmol, 73%).

δ_H (MeOD, 500 MHz): 4.54 (1H, q, *J* 7.0 Hz, H-5), 3.71 (3H, s, H-11), 1.44 (9H, s, H-1), 1.39 (3H, s, H-9), 1.37 (3H, s, H-9), 1.22 (3H, d, *J* 7.0 Hz, H-6).

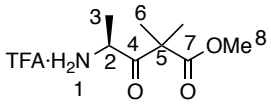
δ_C (MeOD, 125 MHz): 209.5 (C-7), 175.1 (C-10), 157.3 (C-3), 80.6 (C-2), 55.4 (C-8), 53.0 (C-11), 52.8 (C-5), 28.1 (C-1), 22.8 (C-9), 18.1 (C-6).

HRMS (ESI^+): calcd for $C_{13}H_{23}NNaO_5^+$: 296.1468, found: 296.1471.

IR (cm^{-1}): 1748 (C=O stretch), 1713 (C=O stretch), 1677 (C=O stretch).

$[\alpha]_D^{20}$ = -24.7 (*c* 0.255, $CHCl_3$).

158 *(S)*-5-Methoxy-4,4-dimethyl-3,5-dioxopentan-2-ammonium trifluoroacetate



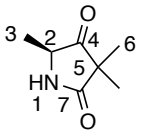
Under an atmosphere of argon to a 10% TFA in DCM solution (5 mL) while stirring was added **144** (105 mg, 386 μmol). After half an hour the solution was diluted with DCM (3 × 15 mL), removing solvent in *vacuo* successively. The resulting crude oil was washed with petroleum ether (3 × 5 mL), and residual solvent removed in *vacuo* from the immiscible oil to give the desired product as a brown oil (107 mg, 373 μmol, 97%).

δ_H (MeOD, 500 MHz): 4.51 (1H, q, *J* 6.5 Hz, H-2), 3.76 (3H, s, H-8), 1.48 (3H, s, H-6), 1.47 (3H, s, H-6), 1.45 (3H, d, *J* 7.0 Hz, H-3).

δ_C (MeOD, 125 MHz): 206.1 (C-4), 174.4 (C-7), 55.4 (C-5), 53.4 (C-8), 53.2 (C-2), 23.0 (C-6), 22.5 (C-6), 16.9 (C-3).

HRMS (ESI^+): calcd for $C_8H_{16}NO_3^+$: 174.1125, found: 174.1126.

159 *(S)*-3,3,5-Trimethylpyrrolidine-2,4-dione



Under an atmosphere of argon to a stirred solution of **93** (50.6 mg, 279 μmol), **158** (79.1 mg, 275 μmol) and HOBt (47.2 mg, 349 μmol) in THF (5.5 mL) at 0 °C was added triethylamine (160 μL, 1.1 mmol) and EDC (105 mg, 549 μmol). The reaction was warmed to room temperature and stirred for 24 hours. Reaction mixture was filtered, and solvent removed in *vacuo*. The resulting residue was purified by silica gel chromatography (ethyl acetate) to provide the lactam byproduct **159** as a white solid (31.4 mg, 222 μmol, 80%) and none of the desired **154**.[104]

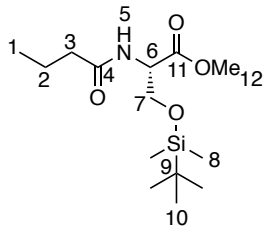
δ_H (MeOD, 500 MHz): 4.11 (1H, q, *J* 7.0 Hz, H-2), 1.32 (3H, d, *J* 7.0 Hz, H-3), 1.20 (3H, s, H-6), 1.18 (3H, s, H-6).

δ_C (MeOD, 125 MHz): 216.3 (C-4), 180.1 (C-7), 64.4 (C-5), 58.1 (C-2), 21.4 (C-6), 21.0 (C-6), 18.0 (C-3).

m/z ES^+ (%): 163.7 (100, $[M+Na]^+$).

7.3.4 Synthesis of alanine based α-dimethyl-β-keto methyl ester substrate analogue

161 Methyl *O*-TBS-*N*-butyryl-L-serinate



Under an argon atmosphere to a stirred solution of **104** (4.4 g, 23.5 mmol) and imidazole (3.9 g, 56.7 mmol) in DMF (75 mL) was added TBS-chloride (4.3 g, 28 mmol) in DMF (50 mL) dropwise over 1 hour. Upon full addition of the TBS-chloride solution, the reaction was stirred for 24 hours. The reaction was diluted with diethyl ether (250 mL), washed with HCl (1 × 250 mL, 1 M), saturated $NaHCO_3$ (1 × 250 mL) and brine (3 × 250 mL). The organic layer was dried, solvent removed in *vacuo*, and the resulting residue purified by silica gel chromatography (3:1 *n*-hexane:ethyl acetate) to give the desired product as a yellow oil (5.8 g, 19.0 mmol, 81%).

δ_H ($CDCl_3$, 500 MHz): 6.27 (1H, br d, *J* 8.0 Hz, H-5), 4.66 (1H, dt, *J* 8.0, 3.0 Hz, H-6), 4.03 (1H, dd, *J* 10.0, 2.5 Hz, H-7), 3.80 (1H, dd, *J* 10.0, 3.0 Hz, H-7), 3.72 (3H, s, H-12), 2.22 (2H, br t, *J* 7.5 Hz, H-3) 1.66 (2H, sextet, *J* 7.5 Hz, H-2), 0.95 (3H, t, *J* 7.5 Hz, H-1), 0.84 (9H, s, H-10), 0.01 (3H, s, H-8), -0.00 (3H, s, H-8).

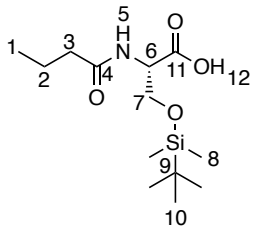
δ_C ($CDCl_3$, 125 MHz): 172.8 (C-4), 171.1 (C-11), 63.7 (C-7), 54.2 (C-6), 52.4 (C-12), 38.6 (C-3), 25.8 (C-10), 19.1 (C-2), 18.2 (C-9), 13.8 (C-1), -5.5 (C-8), -5.6 (C-8).

HRMS (ESI^+): calcd for $C_{14}H_{29}NNaO_4Si^+$: 326.1758, found: 326.1767.

IR (cm^{-1}): 1747 (C=O stretch), 1650 (C=O stretch).

$[\alpha]_D^{20}$ = +36.1 (*c* 0.478, $CHCl_3$).

162 *O*-TBS-*N*-butanoyl-L-serine



161 (4.5 g, 14.8 mmol) was dissolved in a mixture of THF (35 mL) and aqueous lithium hydroxide (30 mL, 0.75 M, 22.4 mmol). The reaction was stirred for half an hour, then THF removed in *vacuo*. The solution was washed with diethyl ether (5 × 50 mL), pH adjusted to 7 with HCl (1 M), water removed in *vacuo* and the resulting residue dissolved in ethyl acetate (40 mL). The solution was dried ($MgSO_4$), filtered and solvent removed in *vacuo* to yield the desired product as a pale pink solid (3.0 g, 10.3 mmol, 69%).

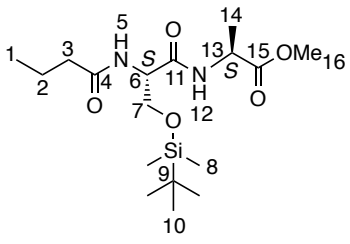
δ_H (MeOD, 500 MHz): 4.33 (1H, t, *J* 4.5 Hz, H-6), 3.97 (2H, d, *J* 4.5 Hz, H-7), 2.24 (2H, t, *J* 7.5 Hz, H-3), 1.66 (2H, sextet, 7.5 Hz, H-2), 0.97 (3H, t, *J* 7.5 Hz, H-1), 0.89 (9H, s, H-10), 0.06 (3H, s, H-8), 0.05 (3H, s, H-8).

δ_C (MeOD, 125 MHz): 177.2 (C-11), 175.4 (C-4), 65.0 (C-7), 58.1 (C-6), 39.2 (C-3), 26.4 (C-10), 20.3 (C-2), 19.2 (C-9), 14.1 (C-1), -5.2 (C-8).

HRMS (ESI^+): calcd for $C_{13}H_{27}NNaO_4Si^+$: 312.1602, found: 312.1604.

IR (cm^{-1}): 1702 (C=O stretch), 1656 (C=O stretch).

163 Methyl *O*-TBS-*N*-butyryl-L-seryl-L-alaninate



Under an atmosphere of argon to a stirred solution of **162** (1.2 g, 4.3 mmol), L-alanine methyl ester hydrochloride salt (597 mg, 4.3 mmol) and HOBt (1.4 g, 10.3 mmol) in THF (50 mL) at 0 °C was added triethylamine (2.5 mL, 17.9 mmol) and EDC (1.6 g,

8.5 mmol). The reaction was warmed to room temperature and stirred for 24 hours. Reaction mixture was filtered, and solvent removed in *vacuo*. The resulting residue was purified by silica gel chromatography (gradient of 1:1 *n*-hexane:ethyl acetate to ethyl acetate) to yield the desired product as yellow solid (1.0 g, 2.8 mmol, 65%). ^{1}H NMR spectroscopic analysis indicated that it contained approximately 20% of the C-6 epimer. The data reported are for the predicted major diastereomer (*SS*).

δ_H (MeOD, 500 MHz): 4.54-4.35 (2H, m, H-6, H-13), 3.91-3.81 (2H, m, H-7), 3.71 (3H, s, H-16), 2.25 (2H, t, *J* 7.5 Hz, H-3), 1.65 (2H, sextet, *J* 7.5 Hz, H-2), 1.38 (3H, d, *J* 7.5 Hz, H-14), 0.96 (3H, t, *J* 7.5 Hz, H-1), 0.90 (9H, s, H-10), 0.09 (6H, s, H-8).

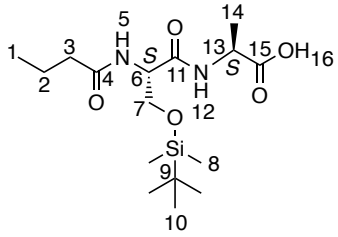
δ_C (MeOD, 125 MHz): 176.1 (C-4), 174.4 (C-16), 171.7 (C-11), 64.3 (C-7), 56.4 (C-6), 52.8 (C-16), 49.9 (C-13), 38.8 (C-3), 26.3 (C-10), 18.8 (C-2), 17.7 (C-14), 16.5 (C-9), 14.0 (C-1),

-5.4 (C-8).

HRMS (ESI$^+$): calcd for $C_{17}H_{34}N_2NaO_5Si^+$: 397.2129, found: 397.2137.

IR (cm^{-1}): 1745 (C=O stretch).

164 *O*-TBS-*N*-butanoyl-L-seryl-L-alanine



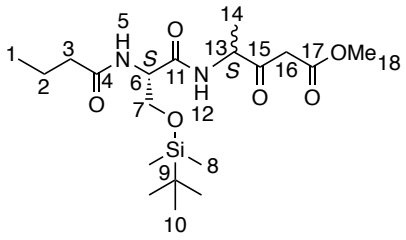
163 (467 mg, 1.2 mmol) was dissolved in a mixture of THF (3 mL) and aqueous lithium hydroxide (2.5 mL, 0.82 M, 1.9 mmol). The reaction was stirred for half an hour, pH adjusted to 1 using HCl (1 M), and THF removed in *vacuo*. Solution was diluted with water (5 mL), sodium chloride added (ca. 0.5 g) and extracted with ethyl acetate (5 × 50 mL). Organic extractions were dried ($MgSO_4$) and solvent removed in *vacuo* to yield the desired product as a clear oil (369 mg, 1.0 mmol, 82%). ^{1}H NMR spectroscopic analysis indicated that it contained approximately 20% of the C-6 epimer. The data reported are for the predicted major diastereomer (*SS*).

δ_H (MeOD, 500 MHz): 4.46 (1H, t, *J* 5.5 Hz, H-6), 4.16 (1H, q, *J* 7.0 Hz, H-13), 3.93-3.85 (1H, m, H-7), 2.35-2.22 (2H, m, H-3), 1.67 (2H, sextet, *J* 7.5 Hz, H-2), 1.35 (3H, d, *J* 7.0 Hz, H-14), 0.97 (3H, t, *J* 7.5 Hz, H-1), 0.90 (9H, s, H-10), 0.09 (3H, s, H-8), 0.08 (3H, s, H-8).

δ_C (MeOD, 125 MHz): 179.0 (C-15), 176.2 (C-4), 170.9 (C-11), 64.3 (C-7), 56.8 (C-6), 52.1 (C-13), 38.9 (C-3), 26.4 (C-10), 20.3 (C-2), 19.7 (C-14), 19.2 (C-9), 14.1 (C-1), -5.3 (C-8).

HRMS (ESI^-): calcd for $C_{16}H_{31}N_2O_5Si^-$: 359.2008, found: 359.2008.

165 Methyl *(S)*-4-*((S)*-3-((TBS)oxy)-2-butyramidopropanamido)-3-oxopentanoate



Under an atmosphere of argon to a stirred solution of **164** (215 mg, 6.0 mmol) in 1,4-dioxane (8.3 mL) at 50 °C was added CDI (118 mg, 7.3 mmol). The solution was stirred at 50 °C for two hours. Methyl potassium malonate (141 mg, 9.0 mmol) and $MgCl_2$ (86.4 mg, 7.2 mmol) were added and the reaction allowed to stir for 24 hours maintaining 50 °C. Solvent was removed in *vacuo*, and the crude residue was partitioned between ethyl acetate (50 mL) and HCl (1 M, 50 mL). The organic layer was collected, and the aqueous phase was extracted with ethyl acetate (5 × 50 mL). Organic extractions were combined, washed with 5% $NaHCO_3$ (*w/w*, 2 × 50 mL) and brine (50 mL), and dried ($MgSO_4$). Solvent was removed in *vacuo* to yield the desired product as a clear oil 3:2 mixture of diastereomers resulting from epimerization at C-13 (196 mg, 4.7 mmol, 79%).

HRMS (ESI^+): calcd for $C_{19}H_{36}N_2NaO_6Si^+$: 439.2235, found: 439.2238.

IR (cm^{-1}): 1739 (C=O stretch), 1723 (C=O stretch).

Predicted major diastereomer (*SS*):

δ_H ($CDCl_3$, 500 MHz): 7.11 (1H, br d, *J* 7.0 Hz, H-12), 6.38-6.28 (1H, m, H-5), 4.64 (1H, br t, *J* 7.0 Hz, H-13), 4.46-4.39 (1H, m, H-6), 4.05 (1H, br dd, *J* 9.5, 4.0 Hz, H-7), 3.74 (3H, br s, H-18), 3.64-3.41 (3H, m, H-7, H-16), 2.23 (2H, br t, *J* 8.0 Hz, H-3), 1.68 (2H, sextet, *J* 7.5 Hz, H-2), 1.38 (3H, d, *J* 7.0 Hz, H-14), 0.96 (3H, t, *J* 7.3 Hz, H-1), 0.90 (9H, s, H-10), 0.12 (3H, s, H-8), 0.11 (3H, s, H-8).

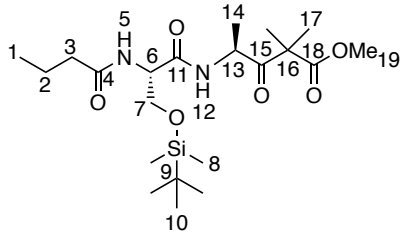
δ_C ($CDCl_3$, 125 MHz): 201.1 (C-15), 173.3 (C-4), 170.4 (C-11), 167.3 (C-17), 62.8 (C-7), 54.6 (C-13), 54.2 (C-6), 52.7 (C-18), 45.8 (C-16), 38.6 (C-3), 25.9 (C-10), 19.1 (C-2), 18.3 (C-9), 16.9 (C-14), 13.9 (C-1), -5.4 (C-8).

Predicted minor diastereomer (*SR*):

δ_H ($CDCl_3$, 500 MHz): 7.28 (1H, br d, *J* 6.6 Hz, H-12), 6.38-6.28 (1H, m, H-5), 4.66 (1H, br t, *J* 7.0 Hz, H-13), 4.46-4.39 (1H, m, H-6), 4.05 (1H, br dd, *J* 9.5, 4.0 Hz, H-7), 3.74 (3H, br s, H-18), 3.64-3.41 (3H, m, H-7, H-16), 2.23 (2H, br t, *J* 8.0 Hz, H-3), 1.68 (2H, sextet, *J* 7.5 Hz, H-2), 1.39 (3H, d, *J* 7.0 Hz, H-14), 0.96 (3H, t, *J* 7.5 Hz, H-1), 0.90 (9H, s, H-10), 0.11 (3H, s, H-8), 0.09 (3H, s, H-8).

δ_C ($CDCl_3$, 125 MHz): 201.1 (C-15), 173.3 (C-4), 170.3 (C-11), 167.2 (C-17), 62.8 (C-7), 54.5 (C-13), 54.2 (C-6), 52.7 (C-18), 45.8 (C-16), 38.6 (C-3), 25.9 (C-10), 19.1 (C-2), 18.3 (C-9), 16.9 (C-14), 13.9 (C-1), -5.3 (C-8), -5.4 (C-8).

166 Methyl *(S)*-4-*((S)*-3-((TBS)oxy)-2-butyramidopropanamido)-2,2-dimethyl-3-oxopentanoate



Under an atmosphere of argon to a stirred solution of **165** (499 mg, 1.2 mmol) was added potassium carbonate (1.7 g, 12.0 mmol) and iodomethane (750 µL, 12.1 mmol). The reaction was stirred at room temperature for 43 hours, filtered and the solvent

removed in *vacuo*. The resulting residue was purified by silica gel chromatography (gradient of 7:3 *n*-hexane:ethyl acetate to 2:3 *n*-hexane:ethyl acetate). The *α*-dimethyl-*β*-keto methyl ester C-13 epimers were separated (*SS* and *SR*), providing the desired predicted major C-13 epimer **166** (*SS*) as a clear oil (147 mg, 331 μmol, 28%) and the predicted minor C-13 epimer **167** (*SR*) as a white powder (79.2 mg, 178 μmol, 15%), as well as a mixture of mono-methylated diastereomers **168** (101 mg, 235 μmol, 20%). The data reported are for the desired major C-13 epimer (*SS*).

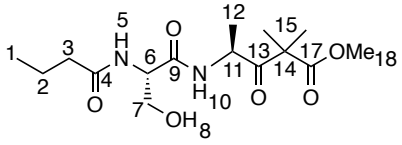
δ_H ($CDCl_3$, 500 MHz): 7.14 (1H, br d, *J* 7.5 Hz, H-12), 6.30 (1H, br d, *J* 6.0 Hz, H-5), 4.97-4.90 (1H, m, H-13), 4.39-4.34 (1H, m, H-6), 4.02 (1H, dd, *J* 9.5, 4.0 Hz, H-7), 3.72 (3H, s, H-19), 3.55 (1H, dd, *J* 9.5, 8.0 Hz, H-7), 2.22 (2H, t, *J* 7.5 Hz, H-3), 1.67 (2H, sextet, *J* 7.5 Hz, H-2), 1.44 (3H, s, H-17), 1.41 (3H, s, H-17), 1.28 (3H, d, *J* 6.9 Hz, H-14), 0.95 (3H, t, *J* 7.5 Hz, H-1), 0.90 (9H, s, H-10), 0.11 (3H, s, H-8), 0.10 (3H, s, H-8).

δ_C ($CDCl_3$, 125 MHz): 207.3 (C-15), 173.5 (C-11), 173.2 (C-4), 169.6 (C-18), 62.8 (C-7), 54.8 (C-16), 54.2 (C-6), 52.8 (C-19), 50.4 (C-13), 38.6 (C-3), 25.9 (C-10), 22.5 (C-17), 22.3 (C-17), 19.1 (C-2), 18.9 (C-14), 18.3 (C-9), 13.9 (C-1), -5.4 (C-8).

HRMS (ESI^+): calcd for $C_{21}H_{40}N_2NaO_6Si^+$: 467.2548, found: 467.2553.

IR (cm^{-1}): 1750 (C=O stretch), 1719 (C=O stretch).

154 Methyl *(S)*-4-*((S)*-2-butyramido-3-hydroxypropanamido)-2,2-dimethyl-3-oxopentanoate



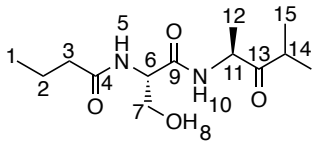
166 (40.1 mg, 90 μmol) was dissolved in an acetic acid, water and THF mixture (3:3:1, 940 μL) and stirred for 17 hours. Toluene was added sequentially (3 × 30 mL) and solvent removed in *vacuo*. Residue was washed with petroleum ether (5 mL) and residual solvent removed in *vacuo* to give the desired product as a cloudy oil (22.7 mg, 69 μmol, 76%).

δ_H (MeOD, 500 MHz): 4.91 (1H, q, *J* 7.0 Hz, H-11), 4.40 (1H, t, *J* 5.5 Hz, H-6), 3.77-3.68 (2H, m, H-7), 3.72 (3H, s, H-18), 2.25 (2H, t, *J* 7.0 Hz, H-3), 1.65 (2H, sextet, 7.0 Hz, H-2), 1.41 (3H, s, H-15), 1.38 (3H, s, H-15), 1.27 (3H, d, *J* 7.0 Hz, H-12), 0.96 (3H, t, *J* 7.5 Hz, H-1).

δ_C (MeOD, 125 MHz): 208.6 (C-13), 176.3 (C-4), 175.0 (C-17), 171.7 (C-9), 63.0 (C-7), 56.7 (C-6), 55.9 (C-14), 53.1 (C-18), 51.3 (C-11), 38.7 (C-3), 22.7 (C-15), 22.6 (C-15), 20.2 (C-2), 18.2 (C-12), 14.0 (C-1).

HRMS (ESI$^+$): calcd for $C_{15}H_{26}N_2NaO_6^+$: 353.1683, found: 353.1680.

171 *N-((S)*-3-Hydroxy-1-*(((S)*-4-methyl-3-oxopentan-2-yl)amino)-1-oxopropan-2-yl)butyramide



154 (11.7 mg, 35 μmol) was dissolved in aqueous lithium hydroxide solution (1.1 mL, 30 mM, 33 μmol) and left to stir. After two hours, **171** byproduct was observed by ESI-MS, and water removed in *vacuo*. The resulting residue was purified by silica gel chromatography (1% methanol in ethyl acetate) to yield the byproduct **171** as a clear oil (2.3 mg, 8.4 μmol, 24%) and allowing for reclamation of **154** (5.3 mg, 16 μmol, 45%). ^{1}H NMR spectroscopic analysis of **171** indicated that it contained approximately 25% of the C-11 epimer. The data reported are for the predicted major diastereomer (*SS*).

δ_H ($CDCl_3$, 500 MHz): 7.07 (1H, br d, *J* 6.5 Hz, H-10), 6.46 (1H, br d, *J* 7.0 Hz, H-5), 4.70 (1H, t, *J* 7.0 Hz, H-11), 4.51-4.46 (1H, m, H-6), 4.08 (1H, dd, *J* 11.5, 3.0 Hz, H-7), 3.77-3.66 (1H, m, H-7), 3.61 (1H, br s, H-8), 2.83 (1H, septet, *J* 7.0 Hz, H-14), 2.23 (2H, t, *J* 7.5 Hz, H-3), 1.68 (2H, sextet, *J* 7.5 Hz, H-2), 1.36 (3H, d, *J* 7.0 Hz, H-12), 1.16 (3H, d, *J* 7.0 Hz, H-15), 1.12 (3H, d, *J* 6.5 Hz, H-15), 0.95 (3H, t, *J* 7.5 Hz, H-1).

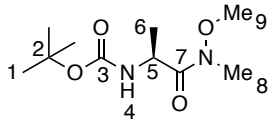
δ_C ($CDCl_3$, 125 MHz): 212.9 (C-13), 173.8 (C-4), 170.8 (C-9), 63.1 (C-7), 53.7 (C-11), 53.00 (C-6), 38.5 (C-3), 37.4 (C-14), 19.2 (C-15), 19.2 (C-2), 17.8 (C-15), 17.2 (C-12), 13.8 (C-1).

HRMS (ESI^+): calcd for $C_{13}H_{24}N_2NaO_4^+$: 295.1628, found: 295.1632.

IR (cm^{-1}): 3253 (O-H stretch), 1715 (C=O stretch).

7.3.4 Synthesis of *α,β*-unsatured ketone authentic standard

187 *tert*-Butyl *(S)*-(1-(methoxy(methyl)amino)-1-oxopropan-2-yl)carbamate



To a stirred solution of *N*-BOC-L-alanine (402 mg, 2.1 mmol) in DCM (9.6 mL) was added O,N-dimethylhydroxylamine hydrochloride (227 mg, 2.3 mmol), triethylamine (620 μL, 4.4 mmol) and BOP reagent (937 mg, 2.1 mmol). Solvent was removed in *vacuo*, and the resulting residue purified by silica gel chromatography (3:2 ethyl acetate:*n*-hexane) to give the desired product as a white solid (266 mg, 1.1 mmol, 54%).[105]

δ_H ($CDCl_3$, 500 MHz): 5.25 (1H, br d, *J* 7.0 Hz, H-4), 4.68 (1H, m, H-5), 3.77 (3H, s, H-9), 3.21 (3H, s, H-8), 1.44 (9H, s, H-1), 1.31 (3H, d, *J* 7.0 Hz, H-6).

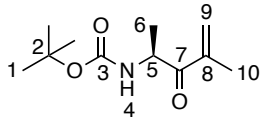
δ_C ($CDCl_3$, 125 MHz): 173.9 (C-7), 155.4 (C-3), 78.0 (C-2), 61.8 (C-9), 46.7 (C-5), 32.3 (C-8), 28.5 (C-1), 18.9 (C-6).

m/z ES^+ (%) 255 (100, $[M+Na]^+$).

IR (cm^{-1}): 1702 (C=O stretch).

$[\alpha]_D^{20}$ = -1.3 (*c* 0.260, $CHCl_3$).

188 *tert*-Butyl *(S)*-(4-methyl-3-oxopent-4-en-2-yl)carbamate



Under an atmosphere of argon to a stirred solution of **187** (200 mg, 863 μmol) in THF (3.5 mL) at 0 °C was added isopropenylmagnesium bromide in THF (500 mM, 3.6 mL, 1.8 mmol) dropwise. The reaction was stirred at 0 °C for 6 hours, then poured into saturated ammonium chloride solution (5 mL) and ice (ca. 5 mL). When ice had melted, pH was adjusted to 7 with HCl (6 M). Solution was extracted with ethyl acetate (3 × 30 mL). Organic extracts were combined and washed with saturated $NaHCO_3$ (2 × 10 mL), brine (2 × 10 mL) and dried ($MgSO_4$). Solvent was removed in *vacuo* to yield the desired product as a pale yellow oil (113 mg, 530 μmol, 61%).[106]

δ_H ($CDCl_3$, 500 MHz): 6.04 (1H, s, H-9), 5.89 (1H, s, H-9), 5.38 (1H, br d, *J* 5.5 Hz, H-4), 5.07-4.99 (1H, m, H-5), 1.91 (3H, s, H-10), 1.44 (9H, s, H-1), 1.31 (3H, d, *J* 7.0 Hz).

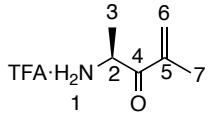
δ_C ($CDCl_3$, 125 MHz): 201.2 (C-7), 155.2 (C-3), 141.9 (C-8), 126.4 (C-9), 79.8 (C-2), 50.3 (C-5), 28.5 (C-1), 20.4 (C-6), 18.0 (C-10).

m/z ES^+ (%) 236 (100, $[M+Na]^+$).

IR (cm^{-1}): 1703 (C=O stretch), 1678 (C=O stretch).

$[\alpha]_D^{20}$ = +38.4 (*c* 0.250, $CHCl_3$).

189 *(S)*-4-methyl-3-oxopent-4-en-2-ammonium trifluoroacetate



To a stirred solution of **188** (97.3 mg, 456 μmol) in DCM (8 mL) was added TFA (2 mL) dropwise. After half an hour the solution was diluted with DCM (3 × 20 mL), removing solvent in *vacuo* successively. The resulting crude oil was washed successively with petroleum ether (3 × 5 mL), and residual solvent removed in *vacuo*

from the immiscible oil to give the desired product as an orange oil (92.7 mg, 408 μmol, 89%).

δ_H (MeOD, 500 MHz): 6.17 (1H, s, H-6), 6.13 (1H, br d, *J* 1.0 Hz, H-6), 4.83 (1H, q, *J* 7.5 Hz, H-2), 1.93 (3H, s, H-7), 1.49, (3H, d, *J* 7.0 Hz, H-3).

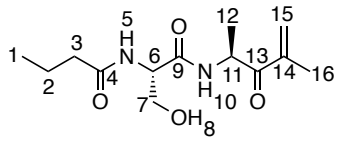
δ_C (MeOD, 125 MHz): 198.5 (C-4), 142.3 (C-5), 129.2 (C-6), 52.0 (C-2), 18.0 (C-3), 17.6 (C-7).

HRMS (ESI^+): calcd for $C_6H_{12}NO^+$:114.0913, found: 114.0913.

IR (cm^{-1}): 2992 (N-H stretch), 1728 (C=O stretch).

$[\alpha]_D^{20} = +17.2$ (*c* 0.250, $CHCl_3$).

186 *N-((S)*-3-hydroxy-1-*(((S)*-4-methyl-3-oxopent-4-en-2-yl)amino)-1-oxopropan-2-yl)butyramide



Under an atmosphere of argon to a stirred solution of HOBt (72.6 mg, 537 μmol) in THF (4.8 mL) at 0 °C was added **93** (39.0 mg, 223 μmol) in THF (390 μL), **189** (50.4 mg, 222 μmol) in THF (500 μL), triethylamine (130 μL, 933 μmol) and EDC (85.2 mg, 444 μmol). The reaction was warmed to room temperature and stirred for 24 hours. Reaction mixture was filtered, and solvent removed in *vacuo*. The resulting residue was purified by silica gel chromatography (gradient of 5% methanol in ethyl acetate to 10% methanol in ethyl acetate) to yield the desired product as a pale yellow solid (28.4 mg, 105 μmol, 47%).

δ_H (MeOD, 500 MHz): 6.12 (1H, s, H-15), 5.94 (1H, br d, *J* 1.5 Hz, H-15), 5.21 (1H, q, *J* 7.0 Hz, H-11), 4.45 (1H, t, *J* 5.5 Hz, H-6), 3.76 (2H, d, *J* 5.5 Hz, H-7), 2.26 (2H, dt, *J* 7.4, 1.5 Hz, H-3), 1.88 (3H, s, H-16), 1.66 (2H, sextet, *J* 7.5 Hz, H-2), 1.32 (3H, d, *J* 7.0 Hz, H-12), 0.96 (3H, t, *J* 7.5 Hz, H-1). 1.41

δ_C (MeOD, 125 MHz): 201.9 (C-13), 176.3 (C-4), 171.9 (C-9), 143.3 (C-14), 126.9 (C-15), 63.1 (C-7), 56.7 (C-6), 50.7 (C-11), 38.7 (C-3), 20.2 (C-2), 18.7 (C-12), 18.0 (C-16), 14.0 (C-1).

HRMS (ESI^+): calcd for $C_{13}H_{22}N_2NaO_4^+$: 293.1472, found: 293.1467.

IR (cm^{-1}): 3273 (O-H stretch), 1703 (C=O stretch), 1678 (C=O stretch).

7.4. In vitro assays

7.4.1 LC elution profile

UHPLC-ESI-Q-TOF MS analyses of supernatant from assays containing **66** or **154** as a substrate, or **186** as an authentic standard, used the following elution profile:

Table 7.5: Elution profile for assays containing either 66 or 154 as a substrate.

Time / minutes	Acetontrile + 0.1% formic acid percentage	Water + 0.1% formic acid percentage
0	5	95
5	5	95
17	100	0
22	100	0
25	5	95
33	5	95

UHPLC-ESI-Q-TOF MS analyses of supernatant from assays containing **139** as a substrate used the following elution profile:

Table 7.6: Elution profile for assays containing 139 as a substrate.

Time / minutes	Acetontrile + 0.1% formic acid percentage	Water + 0.1% formic acid percentage
0	5	95
5	5	95
19	100	0
24	100	0
27	5	95
35	5	95

7.4.2. EpnF activity assay using 66 substrate

EpnF (20 μM) was incubated with **66** (2 mM) in concentration buffer (200 μL final volume, section 7.2.6) at 25 °C for 3 hours. The reaction was stopped and protein precipitated by addition of 200 μL acetonitrile containing 0.1% formic acid. The precipitate was pelleted by centrifugation and the supernatant was analysed by UHPLC-ESI-Q-TOF MS.[42]

7.4.2.1 Trapping of α,β-unsaturated ketone intermediate in EpnF-catalysed reaction

N-Acetylcysteamine (total concentration 4 mM) was added to the EpnF catalysed reaction. UHPLC-ESI-Q-TOF MS analysis showed that the both products **136** and **137** formed.[42]

7.4.3 TmcI activity assay with EpnF using 66 substrate

EpnF (10 μM) and TmcI (10 μM) were incubated with **66** (2 mM), ferredoxin (1 μM) and ferredoxin-$NADP^+$ reductase (1 μM) from *Spinacia oleracea*, and NADPH (4 mM) in concentration buffer (200 μL final volume, section 7.2.6) at 25 °C for 3 hours. The reaction was stopped by addition of 200 μL acetonitrile containing 0.1% formic acid. The precipitated protein was pelleted by centrifugation and the supernatant was analysed by high resolution LC-MS/MS.[42]

7.4.4 Pig liver esterase activity assay

Pig liver esterase (10 μM) was incubated with either **139** or **154** (1 mM) in concentration buffer (200 μL final volume, pH 8, section 7.2.6) at 25 °C for 3 hours. The reaction was stopped and protein precipitated by addition of 200 μL acetonitrile containing 0.1% formic acid. The precipitate was pelleted by centrifugation and the supernatant was analysed by UHPLC-ESI-Q-TOF MS.

7.4.5 EpnF and pig liver esterase tandem assay

EpnF (10 μM) and pig liver esterase (10 μM) were incubated with either **139** or **154** (1 mM) in concentration buffer (200 μL final volume, pH 8, section 7.2.6) at 25 °C for 3 hours. The reaction was stopped and protein was precipitated by by addition of 200 μL acetonitrile containing 0.1% formic acid. The precipitate was pelleted by centrifugation and the supernatant was analysed by UHPLC-ESI-Q-TOF MS in the case of **139** and LC-MS/MS in the case of **154**. **186** was co-injected with the supernatant of a **154** assay as an authentic standard for a substrate intermediate.

7.4.6 EpnF, PLE and FAD or FMN assay

To **154** assays set up as per section 7.4.5 was added either FAD or FMN (2 mM). The precipitated protein was pelleted by centrifugation and the supernatant was analysed by UHPLC-ESI-Q-TOF MS.

7.4.7 EpnF nicotinamide adenine dinucleotide reduction

EpnF (20 μM) in Tris-HCl pH 8 buffer (25 mM) was mixed with varying concentrations of NADH (20 – 500 μM final concentration) and absorbance measured at 340 nm over 10 minutes. A decrease in absorbance was observed in each case, corresponding to the conversion of NADH to NAD^+.

References

(1) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2007**, *70*, 461.

(2) Dayan, F. E.; Cantrell, C. L.; Duke, S. O. *Bioorganic Med. Chem.* **2009**, *17*, 4022.

(3) Wani, M. C.; Taylor, H. L.; Wall, M. E.; Coggon, P.; Mcphail, A. T. *J. Am. Chem. Soc.* **1971**, *93*, 2325.

(4) Croteau, R.; Ketchum, R. E. B.; Long, R. M.; Kaspera, R.; Wildung, M. R. *Phytochem. Rev.* **2006**, *5*, 75.

(5) Ajikumar, P. K.; Tyo, K.; Carlsen, S.; Mucha, O.; Phon, T. H.; Stephanopoulos, G. *Mol. Pharm*. **2008**, *5*, 167.

(6) Weber, J. M.; Leung, J. O.; Swanson, S. J.; Idler, K. B.; McAlpine, J. B. *Science* **1991**, *252*, 114.

(7) Zhanel, G. G.; Dueck, M.; Hoban, D. J.; Vercaigne, L. M.; Embil, J. M.; Gin, A. S.; Karlowsky, J. A. *Drugs* **2001**, *61*, 443.

(8) Xu, Y.; Orozco, R.; Kithsiri Wijeratne, E. M.; Espinosa-Artiles, P.; Leslie Gunatilaka, A. A.; Patricia Stock, S.; Molnár, I. *Fungal Genet. Biol.* **2009**, *46*, 353.

(9) Nett, M.; Ikeda, H.; Moore, B. S. *Nat. Prod. Rep.* **2009**, *26*, 1362.

(10) Flärdh, K.; Buttner, M. J. *Nat. Rev. Microbiol.* **2009**, *7*, 36.

(11) Watve, M. G.; Tickoo, R.; Jog, M. M.; Bhole, B. D. *Arch. Microbiol*. **2001**, *176*, 386.

(12) Schatz, A.; Bugle, E.; Waksman, S. A. *Exp. Biol. Med.* **1944**, *55*, 66.

(13) Zumla, A.; Nahid, P.; Cole, S. T. *Nat. Rev. Drug Discov.* **2013**, *12*, 388.

(14) Noomnual, S.; Thasana, N.; Sungkeeree, P.; Mongkolsuk, S.; Loprasert, S. *J. Antibiot. (Tokyo)*. **2015**, *69*, 1.

(15) Du, L.; Sánchez, C.; Chen, M.; Edwards, D. J.; Shen, B. *Chem. Biol.* **2000**, *7*, 623.

(16) Giloni, L.; Takeshita, M.; Johnson, F.; Iden, C.; Grollman, A. P. *J. Biol. Chem.* **1981**, *256*, 8608.

(17) Kim, K. B.; Crews, C. M. *Nat. Prod. Rep.* **2013**, *30*, 600.

(18) Sugawara, K.; Hatori, M.; Nishiyama, Y.; Tomita, K.; Kamei, H.; Konishi, M.; Oki, T. *J. Antibiot. (Tokyo)*. **1990**, *43*, 8.

(19) Hanada, M.; Sugawara, K.; Kaneta, K.; Toda, S.; Nishiyama, Y.; Tomita, K.; Yamamoto, H.; Konishi, M.; Oki, T. *J. Antibiot. (Tokyo)*. **1992**, *45*, 1746.

(20) Tsuchiya, K.; Kobayashi, S.; Nishikiori, T.; Nakagawa, T.; Tatsuta, K. *J. Antibiot. (Tokyo)*. **1997**, *50*, 261.

(21) Koguchi, Y.; Kohno, J.; Suzuki, S.; Nishio, M.; Takahashi, K.; Ohnuki, T.; Komatsubara, S. *J. Antibiot. (Tokyo)*. **1999**, *52*, 1069.

(22) Pereira, A. R.; Kale, A. J.; Fenley, A. T.; Byrum, T.; Debonsi, H. M.; Gilson, M. K.; Valeriote, F. A.; Moore, B. S.; Gerwick, W. H. *Chembiochem* **2012**, *13*, 810.

(23) Tsunematsu, Y.; Nishimura, S.; Hattori, A.; Oishi, S.; Fujii, N.; Kakeya, H. *Org. Lett.* **2015**, *17*, 258.

(24) Owen, J. G.; Charlop-Powers, Z.; Smith, A. G.; Ternei, M. a; Calle, P. Y.; Reddy, B. V. B.; Montiel, D.; Brady, S. F. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 4221.

(25) Keller, L.; Plaza, A.; Dubiella, C.; Groll, M.; Kaiser, M.; Müller, R. *J. Am. Chem. Soc.* **2015**, *137*, 8121.

(26) Rui, Z.; Petříčková, K.; Škanta, F.; Pospíšil, S.; Yang, Y.; Chen, C. Y.; Tsai, S. F.; Floss, H. G.; Petříček, M.; Yu, T. W. *J. Biol. Chem.* **2010**, *285*, 24915.

(27) Zeeck, A.; Schroder, K.; Frobel, K.; Grote, R.; Thiericke, R. *J Antibiot* **1987**, *40*, 1530.

(28) Kijima, M.; Yoshida, M.; Sugita, K.; Horinouchi, S.; Beppu, T. *J. Biol. Chem.* **1993**, *268*, 22429.

(29) Porter, N. J.; Christianson, D. W. *ACS Chem. Biol.* **2017**, *12*, 2281.

(30) Lander, G. C.; Estrin, E.; Matyskiela, M. E.; Bashore, C.; Nogales, E.; Martin, A. *Nature*. **2012**, *482*, 186.

(31) Lecker, S. H.; Goldberg, A. L.; Mitch, W. E. *J. Am. Soc. Nephrol.* **2006**, *17*, 1807.

(32) Da Fonseca, P. C. A.; Morris, E. P. *J. Biol. Chem.* **2008**, *283*, 23305.

(33) Meng, L.; Kwok, B. H. B.; Sin, N.; Crews, C. M. *Cancer Res.* **1999**, *59*, 2798.

(34) Schrader, J.; Henneberg, F.; Mata, R. A.; Tittmann, K.; Schneider, T. R.; Stark, H.; Bourenkov, G.; Chari, A. *Science* **2016**, *353*, 594.

(35) Groll, M.; Kim, K. B.; Kairies, N.; Huber, R.; Crews, C. M. *J. Am. Chem. Soc.* **2000**, *122*, 1237.

(36) Baldwin, J. E. *J. Chem. Soc. Chem. Commun.* **1976**, *18*, 734.

(37) Carmony, K.; Lee, W.; Kim, K. B. *ChemBioChem* **2016**, *17*, 2115.

(38) Elofsson, M.; Splittgerber, U.; Myung, J.; Mohan, R.; Crews, C. M. *Chem. Biol.* **1999**, *6*, 811.

(39) Lawasut, P.; Chauhan, D.; Laubach, J.; Hayes, C.; Fabre, C.; Maglio, M.; Mitsiades, C.; Hideshima, T.; Anderson, K. C.; Richardson, P. G. *Curr. Hematol. Malig. Rep.* **2012**, *7*, 258.

(40) Sin, N.; Kyung, B. K.; Elofsson, M.; Meng, L.; Auth, H.; Kwok, B. H. B.; Crews, C. M. *Bioorganic Med. Chem. Lett.* **1999**, *9*, 2283.

(41) Schorn, M.; Zettler, J.; Noel, J. P.; Dorrestein, P. C.; Moore, B. S.; Kaysser, L. *ACS Chem. Biol.* **2014**, *9*, 301.

(42) Zabala, D.; Cartwright, J. W.; Roberts, D. M.; Law, B. J. C.; Song, L.; Samborskyy, M.; Leadlay, P. F.; Micklefield, J.; Challis, G. L. *J. Am. Chem. Soc.* **2016**, *138*, 4342.

(43) Li, Y.; Florova, G.; Reynolds, K. A. *J. Bacteriol.* **2005**, *187*, 3795.

(44) Finking, R.; Marahiel, M. A. *Annu. Rev. Microbiol.* **2004**, *58*, 453.

(45) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380.

(46) Santos, C. L.; Correia-Neves, M.; Moradas-Ferreira, P.; Mendes, M. V. *PLoS One* **2012**, *7*, e46758.

(47) Kale, A. J.; McGlinchey, R. P.; Lechner, A.; Moore, B. S. *ACS Chem. Biol.* **2011**, *6*, 1257.

(48) Davies, C.; Heath, R. J.; White, S. W.; Rock, C. O. *Structure* **2000**, *8*, 185.

(49) Lai, C. Y.; Cronan, J. E. *J. Bacteriol.* **2004**, *186*, 1869.

(50) Finzel, K.; Nguyen, C.; Jackson, D. R.; Gupta, A.; Tsai, S. C.; Burkart, M. D.

Chem. Biol. **2015**, *22*, 1453.

(51) Heath, R. J.; Su, N.; Murphy, C. K.; Rock, C. O. *J. Biol. Chem.* **2000**, *275*, 40128.

(52) Challis, G. L.; Ravel, J.; Townsend, C. A. *Chem. Biol.* **2000**, *7*, 211.

(53) Keatinge-Clay, A. T. *Nat. Prod. Rep.* **2012**, *29*, 1050.

(54) Bach, R. D.; Canepa, C. *J. Org. Chem.* **1996**, *61*, 6346.

(55) Westheimer, F. H.; Jones, W. A. *J. Am. Chem. Soc.* **1941**, *63*, 3283.

(56) Ghisla, S.; Thorpe, C. *Eur. J. of Biochem*. **2004**, *271*, 494.

(57) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 171.

(58) Hu, D. X.; Withall, D. M.; Challis, G. L.; Thomson, R. J. *Chem. Rev*. **2016**, *116*, 7818.

(59) Thibodeaux, C. J.; Chang, W.; Liu, H. *Chem. Rev.* **2012**, *112*, 1681.

(60) Groves, J. T. *J. Inorg. Biochem.* **2006**, *100*, 434.

(61) Chun, Y. J.; Shimada, T.; Sanchez-Ponce, R.; Martin, M. V.; Lei, L.; Zhao, B.; Kelly, S. L.; Waterman, M. R.; Lamb, D. C.; Guengerich, F. P. *J. Biol. Chem.* **2007**, *282*, 17486.

(62) Rittle, J.; Green, M. T. *Science.* **2010**, *330*, 933.

(63) Ogura, H.; Nishida, C. R.; Hoch, U. R.; Perera, R.; Dawson, J. H.; Ortiz De Montellano, P. R. *Biochemistry* **2004**, *43*, 14712.

(64) Anzai, Y.; Li, S.; Chaulagain, M. R.; Kinoshita, K.; Kato, F.; Montgomery, J.; Sherman, D. H. *Chem. Biol.* **2008**, *15*, 950.

(65) Lutz, R.; Bujard, H. *Nucleic Acids Res.* **1997**, *25*, 1203.

(66) Studier, F. W.; Rosenberg, A. H.; Dunn, J. J.; Dubendorff, J. W. *Methods Enzymol.* **1990**, *185*, 60.

(67) Dodani, S. C.; Cahn, J. K. B.; Heinisch, T.; Brinkmann-Chen, S.; McIntosh, J. A.; Arnold, F. H. *ChemBioChem.* **2014**, *15*, 2259.

(68) Rosano, G. L.; Ceccarelli, E. A. *Front.Microbiol*. **2014**, *5*, 172.

(69) Qing, G.; Ma, L.-C.; Khorchid, A.; Swapna, G. V. T.; Mal, T. K.; Takayama, M. M.; Xia, B.; Phadtare, S.; Ke, H.; Acton, T.; Montelione, G. T.; Ikura, M.; Inouye, M. *Nat. Biotechnol.* **2004**, *22*, 877.

(70) Rahmanpour, R.; King, L. D. W.; Bugg, T. D. H. *Biochem. Biophys. Res. Commun.* **2017**, *482*, 57.

(71) Denisov, I. G.; Sligar, S. G. *Arch. Biochem. Biophys*. **2012**, *519*, 91.

(72) Barry, S. M.; Challis, G. L. *Methods Enzymol.* **2012**, *516*, 171.

(73) Aliverti, a; Curti, B.; Vanoni, M. a. *Methods Mol. Biol.* **1999**, *131*, 9.

(74) Nahm, S.; Weinreb, S. M. *Tetrahedron Lett.* **1981**, *22*, 3815.

(75) Schmidt, U.; Schmidt, J. *Synthesis (Stuttg.).* **1994**, *3*, 300.

(76) Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron*. **2005**, *3*, 10827.

(77) Oikawa, Y.; Sugano, K; Yonemitsu, O. *J. Org. Chem.* **1978**, *26*, 2087.

(78) Albert, B. J.; Koide, K. *ChemBioChem* **2007**, *8*, 1912.

(79) Liu, J.; Zhu, X.; Zhang, W. *ChemBioChem*. **2015**, *16,* 2585.

(80) Lovey, A. J.; Stephens, W. P.; Krapcho, P. a; Weimaster, J. F.; Eldridge, J. M.; Jahngen, E. G. G. J. *J. Org. Chem.* **1978**, *43*, 138.

(81) Lange, S.; Musidlowska, A.; Schmidt-Dannert, C.; Schmitt, J.; Bornscheuer, U. T. *ChemBioChem* **2001**, *2*, 576.

(82) Domínguez de María, P.; García-Burgos, C.; Bargeman, G.; van Gemert, R. *Synthesis (Stuttg).* **2007**, *10*, 1439.

(83) Björnestedt, R.; Zhong, G.; Lerner, R. A.; Barbas, C. F. *J. Am. Chem. Soc.* **1996**, *118*, 11720.

(84) Jouin, P.; Castro, B.; Nisato, D. *J. Chem. Soc. Perkin Trans. 1* **1987**, *1*, 1177.

(85) Brooks, D. W.; Lu, L. D. L; Masamune, S. *Angew. Chemie Int. Ed. English* **1979**, *18*, 72.

(86) Mansour, T. S.; Evans, C. A. *Synth. Commun.* **1990**, *20*, 773.

(87) Beesley, R. M.; Ingold, C. K.; Thorpe, J. F. *J. Chem. Soc. Trans.* **1915**, *107*, 1080.

(88) Levine, M. N.; Raines, R. T.; Thanassi, J. W.; Cohen, L. A.; Milstien, S.; Cohen, L. A.; Milstien, S.; Cohen, L. A.; Borchardt, R. T.; Cohen, L. A.; Borchardt, R. T.; Cohen, L. A.; Winans, R. E.; Wilcox, C. F.; Danforth, C.; Nicholson, A. W.; James, J. C.; Loudon, G. M.; Wang, B.; Nicolaou, M. G.; Liu, S.; Borchardt, R. T.; Amsberry, K. L.; Gerstenberger, A. E.; Borchardt, R.

T.; Karaman, R.; Bruice, T. C.; Pandit, U. K.; Bryant, R. A. R.; Hansen, D. E.; Radzicka, A.; Wolfenden, R.; Smith, R. M.; Hansen, D. E.; Vyas, D. M.; Wong, H.; Crosswell, A. R.; Casazza, A. M.; Knipe, J. O.; Mamber, S. W.; Doyle, T. W.; Ueda, Y.; Mikkilineni, A. B.; Knipe, J. O.; Rose, W. C.; Casazza, A. M.; Vyas, D. M.; Nicolaou, M. G.; Yuan, C.; Borchardt, R. T.; Nam, N.-H.; Kim, Y.; You, Y.-J.; Hong, D.-H.; Kim, H.-M.; Ahn, B.-Z.; Weerapreeyakul, N.; Anorach, R.; Khuansawad, T.; Yenjai, C.; Isaka, M.; Wani, M. C.; Taylor, H. L.; Wall, M. E.; Coggon, P.; McPhail, A. T.; Dordunoo, S. K.; Burt, H. M.; Rowinsky, E. K.; Cazenave, L. A.; Donehower, R. C.; Amsberry, K. L.; Borchardt, R. T.; Nicolaou, M. G.; Wolfe, J. L.; Schowen, R. L.; Borchardt, R. T.; Gharat, L.; Taneja, R.; Weerapreeyakul, N.; Rege, B.; Polli, J.; Chikhale, P. J.; Lee, H. H.; Palmer, B. D.; Wilson, W. R.; Denny, W. A.; Naughton, D. P.; Dillon, M. P.; Cai, H.; Maag, H.; Harris, J. M.; Chess, R. B.; Maeda, H.; Sawa, T.; Konno, T.; Greenwald, R. B.; Choe, Y. H.; Conover, C. D.; Shum, K.; Wu, D.; Royzen, M.; Wang, B. H.; Gangwar, S.; Pauletti, G. M.; Siahaan, T. J.; Borchardt, R. T.; Zheng, A.; Shan, D.; Wang, B.; Shan, D.; Zheng, A.; Ballard, C. E.; Wang, W.; Borchardt, R. T.; Wang, B.; Cheruvallath, Z. S.; Cole, D. L.; Ravikumar, V. T.; Lavis, L. D.; Raines, R. T.; Lavis, L. D.; Rutkoski, T. J.; Raines, R. T.; Beija, M.; Afonso, C. A. M.; Martinho, J. M. G.; Chandran, S. S.; Dickson, K. A.; Raines, R. T.; Leytus, S. P.; Melhado, L. L.; Mangel, W. F.; Grimm, J. B.; Lavis, L. D.; Lavis, L. D.; Chao, T.-Y.; Raines, R. T.; Watkins, R. W.; Lavis, L. D.; Kung, V. M.; Los, G. V.; Raines, R. T.; Lavis, L. D.; Chao, T.-Y.; Raines, R. T.; Levine, M. N.; Lavis, L. D.; Raines, R. T.; Turcotte, R. F.; Lavis, L. D.; Raines, R. T.; Chao, T.-Y.; Lavis, L. D.; Raines, R. T.; Mangold, S. L.; Carpenter, R. T.; Kiessling, L. L.; Huang, S.-T.; Lin, Y.-L.; Huang, S.-T.; Peng, Y.-X.; Wang, K.-L.; Yatzeck, M. M.; Lavis, L. D.; Chao, T.-Y.; Chandran, S. S.; Raines, R. T.; Zhou, W.; Andrews, C.; Liu, J.; Shultz, J. W.; Valley, M. P.; Cali, J. J.; Hawkins, E. M.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V.; Levine, M. N.; Raines, R. T.; Farinas, E. T.; Schwaneberg, U.; Glieder, A.; Arnold, F. H.; Porstmann, B.; Porstmann, T.; Nugel, E.; Evers,

U.; Cleland, W. W.; Hengge, A. C.; O'Brien, P. J.; Herschlag, D.; Shigenaga, A.; Tsuji, D.; Nishioka, N.; Tsuda, S.; Itoh, K.; Otaka, A.; Shigenaga, A.; Yamamoto, J.; Sumikawa, Y.; Furuta, T.; Otaka, A.; Shigenaga, A.; Yamamoto, J.; Hirakawa, H.; Yamaguchi, K.; Otaka, A.; Shigenaga, A.; Yamamoto, J.; Nishioka, N.; Otaka, A.; Shigenaga, A.; Yamamoto, J.; Hirakawa, H.; Ogura, K.; Maeda, N.; Morishita, K.; Otaka, A.; Siegel, D. P.; Epand, R. M.; Ong, W.; Yang, Y.; Cruciano, A. C.; McCarley, R. L. *Chem. Sci.* **2012**, *3*, 2412.

(89) Patiño-Molina, R.; Martín-Martínez, M.; Herranz, R.; García-López, M. T.; González-Muñiz, R. *Lett. Pept. Sci.* **2000**, *7*, 143.

(90) Moyer, M.; Shiurba, J.; Rapoport, H. *J. Org. Chem.* **1986**, *94720*, 5106.

(91) Kudo, F.; Kitayama, T.; Kakinuma, K.; Eguchi, T. *Tetrahedron Lett.* **2006**, *47*, 1529.

(92) Walsh, C. T.; Chen, Y. -C J. *Angew. Chemie Int. Ed.* **1988**, *27*, 333.

(93) Kantz, A.; Gassner, G. T. *Biochemistry* **2011**, *50*, 523.

(94) Rial, D. V.; Bianchi, D. A.; Kapitanova, P.; Lengar, A.; Van Beilen, J. B.; Mihovilovic, M. D. *European J. Org. Chem.* **2008**, *7*, 1203.

(95) Colonna, S.; Gaggero, N.; Carrea, G.; Ottolina, G.; Pasta, P.; Zambianchi, F. *Tetrahedron Lett.* **2002**, *43*, 1797.

(96) Dawson, R. M. C.; Elliott, D. C.; Elliott, W. H.; Jones, K. M. *Data for biochemical research*; Headington Hill Hall, 1986; Vol. 15.

(97) Wang, Y. W.; Chen, T. Y.; Yang, T. H.; Chang, C. C.; Yang, T. L.; Lo, Y. H.; Huang, J. J. *PLoS One* **2016**, *11*, 169094.

(98) Wang, Z.; Dove, P.; Wang, X.; Shamas-Din, A.; Li, Z.; Nachman, A.; Oh, Y. J.; Hurren, R.; Ruschak, A.; Climie, S.; Press, B.; Griffin, C.; Undzys, E.; Aman, A.; Al-awar, R.; Kay, L. E.; O'Neill, D.; Trudel, S.; Slassi, M.; Schimmer, A. D. *Cell Death Dis.* **2015**, *6*, 1815.

(99) Leroy, E.; Bensel, N.; Reymond, J. L. *Adv. Synth. Catal.* **2003**, *345*, 859.

(100) Romero, P. A.; Arnold, F. H. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866.

(101) Trivella, D. B. B.; Pereira, A. R.; Stein, M. L.; Kasai, Y.; Byrum, T.; Valeriote,

F. A.; Tantillo, D. J.; Groll, M.; Gerwick, W. H.; Moore, B. S. *Chem. Biol.* **2014**, *21*, 782.

(102) Pettit, G. R.; Smith, T. H.; Feng, S.; Knight, J. C.; Tan, R.; Pettit, R. K.; Hinrichs, P. A. *J. Nat. Prod.* **2007**, *70*, 1073.

(103) Theberge, C. R.; Zercher, C. K. *Tetrahedron* **2003**, *59*, 1521.

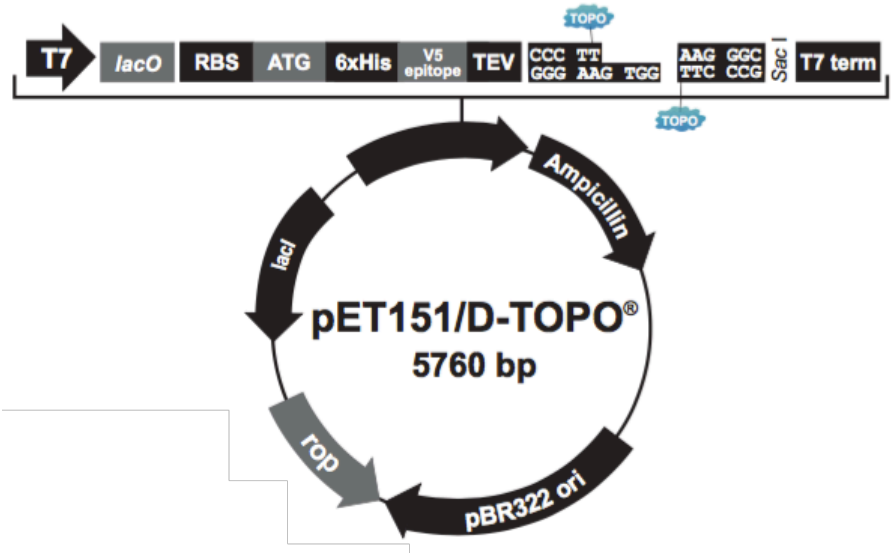
(104) Aikawa, K.; Asano, M.; Ono, K.; Habuka, N.; Yano, J.; Wilson, K.; Fujita, H.; Kandori, H.; Hara, T.; Morimoto, M.; Santou, T.; Yamaoka, M.; Nakayama, M.; Hasuoka, A. *Bioorganic Med. Chem.* **2017**, *25*, 3330.

(105) De Rosa, M.; Lu, L.; Zamaratski, E.; Szałaj, N.; Cao, S.; Wadensten, H.; Lenhammar, L.; Gising, J.; Roos, A. K.; Huseby, D. L.; Larsson, R.; Andrén, P. E.; Hughes, D.; Brandt, P.; Mowbray, S. L.; Karlén, A. *Bioorg. Med. Chem.* **2017**, *25*, 897.

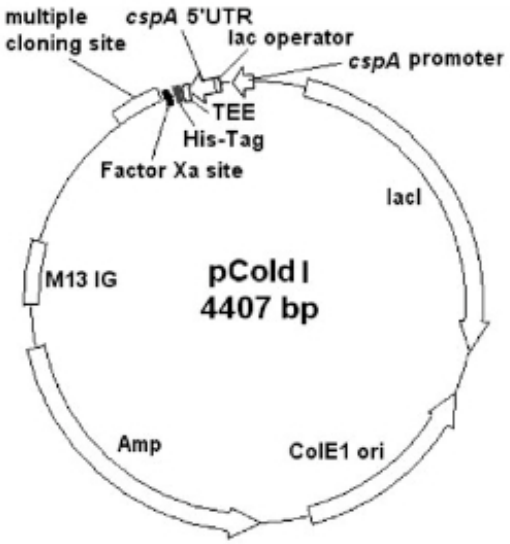
(106) Huber, E. M.; de Bruin, G.; Heinemeyer, W.; Paniagua Soriano, G.; Overkleeft, H. S.; Groll, M. *J. Am. Chem. Soc.* **2015**, *137*, 7835.

Appendix I: Plasmid maps

pET151 (https://www.thermofisher.com/order/catalog/product/K15101)

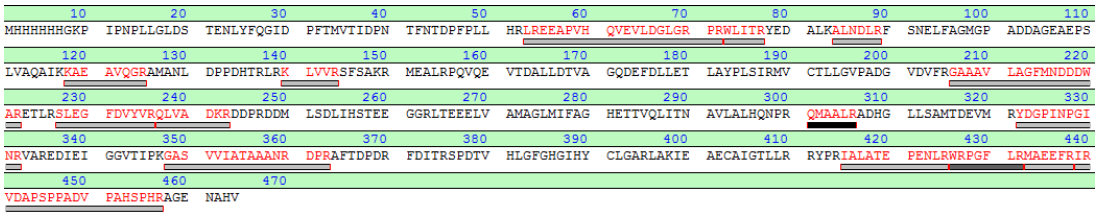


pCold-I (https://www.addgene.org/vector-database/2261/)



Appendix II: Peptide mass fingerprinting

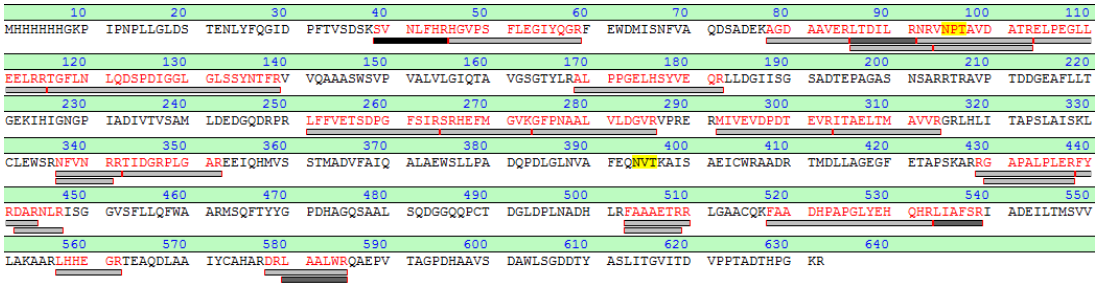
TmcI:



Intensity Coverage: 49.3 % (20029 counts)

Sequence Coverage MS: 33.2%

EpnF:



Intensity Coverage: 69.6 % (121262 counts)

Sequence Coverage MS: 37.2%